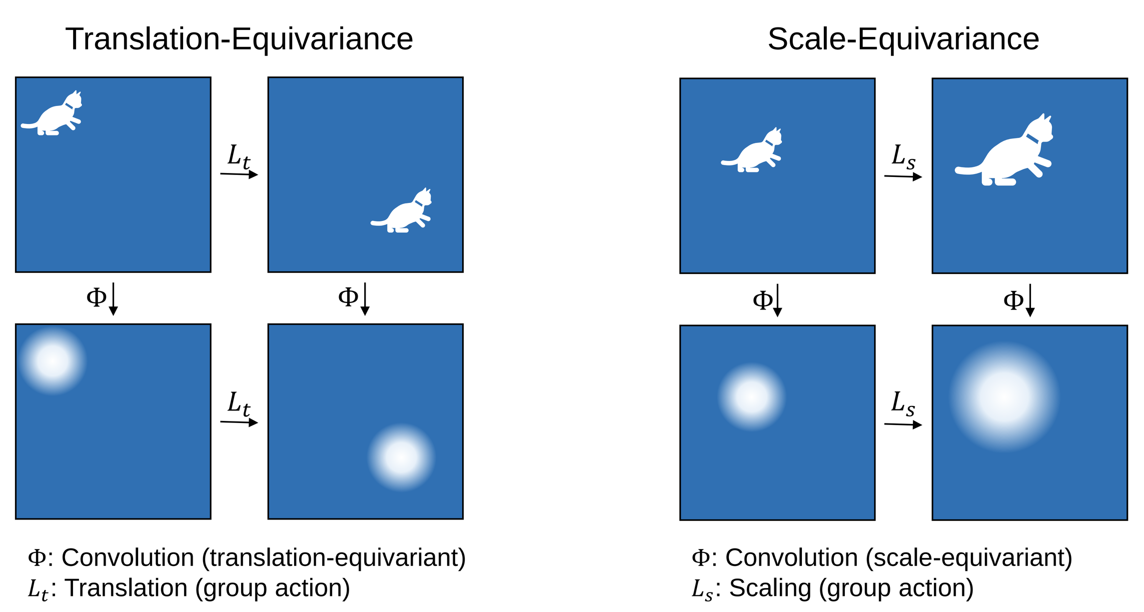
Scale-Equivariant Deep Learning for 3D Data

Thomas Wimmer[1], Vladimir Golkov[1,2], Hoai Nam Dang[3], Moritz Zaiss[3], Andreas Maier[3], Daniel Cremers[1,2]
[1]Technical University of Munich, [2]Munich Center for Machine Learning, [3]Friedrich-Alexander University Erlangen-Nuremberg

Motivation

- Equivariance to group actions ensures that features are detected regardless of their position, rotation, or size.
- Convolutional neural networks (CNNs) are naturally translation-equivariant. We extend this equivariance to scaling of the input data.
- A common training objective is to learn to approximate equivariance through data augmentation. Resulting methods do not guarantee equivariance and handling augmented training data poses an additional learning burden onto the network.
- Equivariant CNN layers give a mathematical guarantee for equivariance and are more data-efficient.
- Medical data is often available at various resolutions (from different scanners), and methods to detect rare diseases need to operate in a low-data regime.



Methods

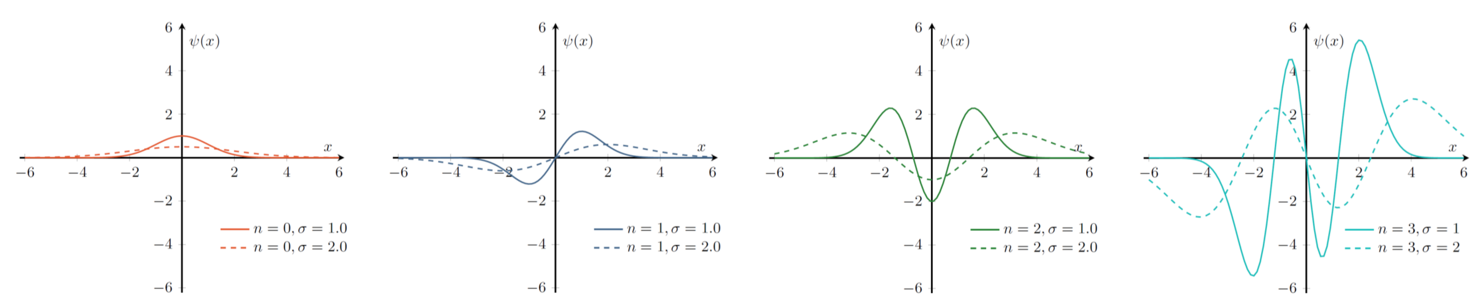
We utilize group-convolutions with the actions of the *group of scalings and translations* $HT = \{(s,t) \mid s \in H, t \in T\}$:

$$L_{(s,t)}[f](x) = f\left(s^{-1}(x - t)\right)$$

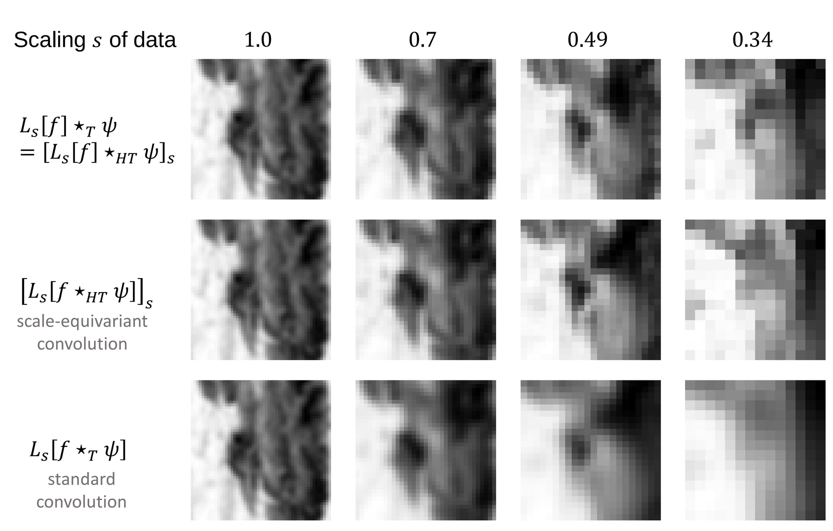
$$L_{(s,t)}[h](s', t') = h\left((s,t)^{-1}(s', t')\right) = h\left(s^{-1}s', s^{-1}(t' - t)\right)$$

on input images $f\colon \mathbb{R}^d \to \mathbb{R}^C$ and latent feature maps $h\colon HT \to \mathbb{R}^C$, respectively.

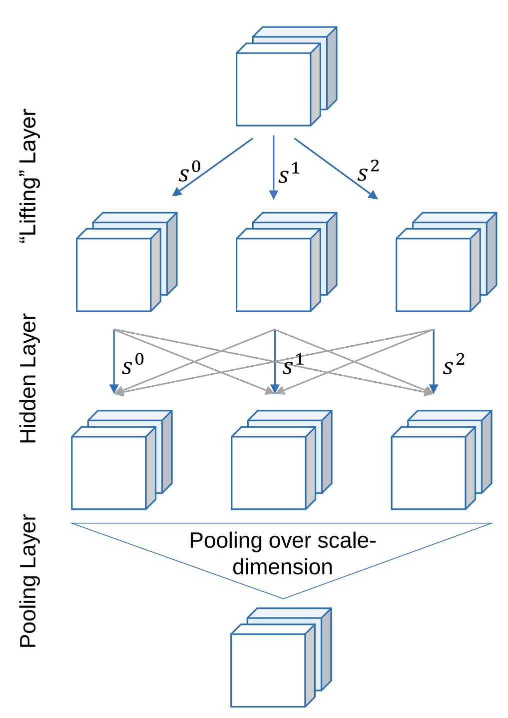
The kernel basis for three-dimensional scale-equivariant convolutions is formed from the multiplication of three oriented basis functions (oriented in the x-, y-, and z-directions) with equal or different degrees of Hermite polynomials.



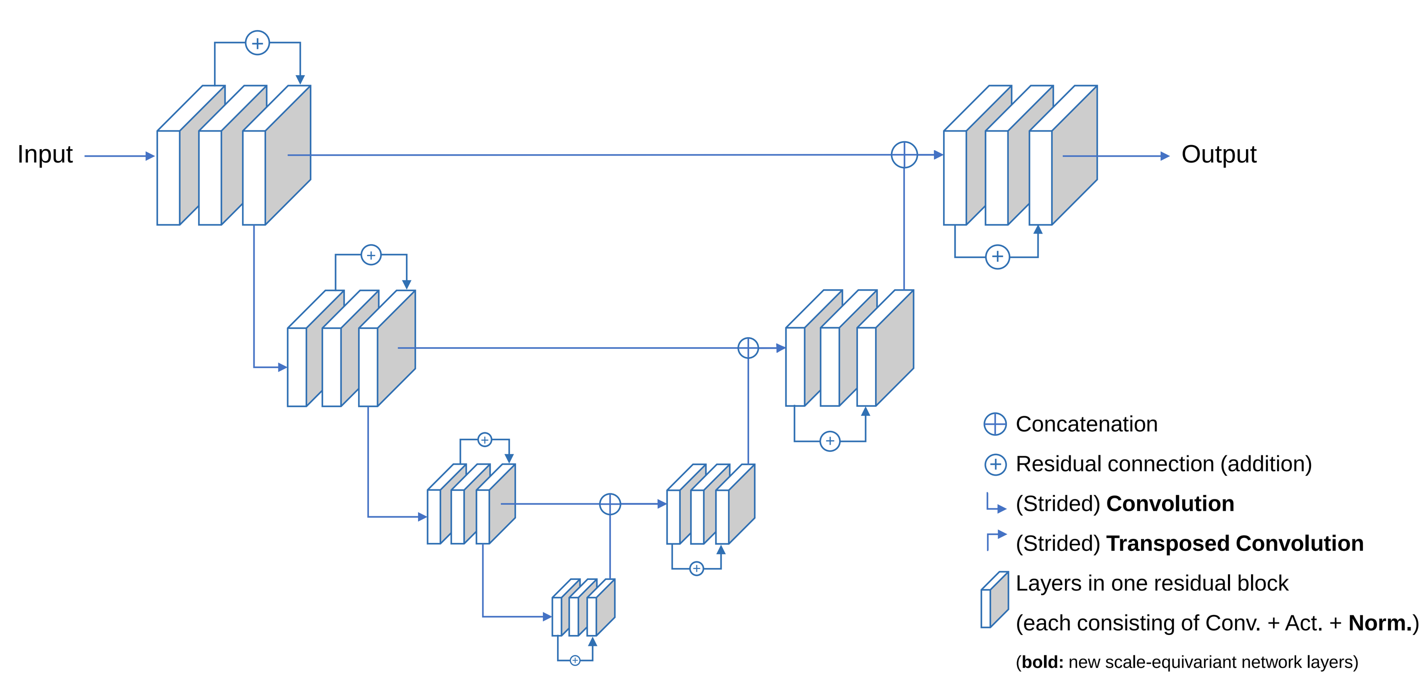
Visualization of equivariance:



Equivariant layer types:



Network Architecture:



Related Work

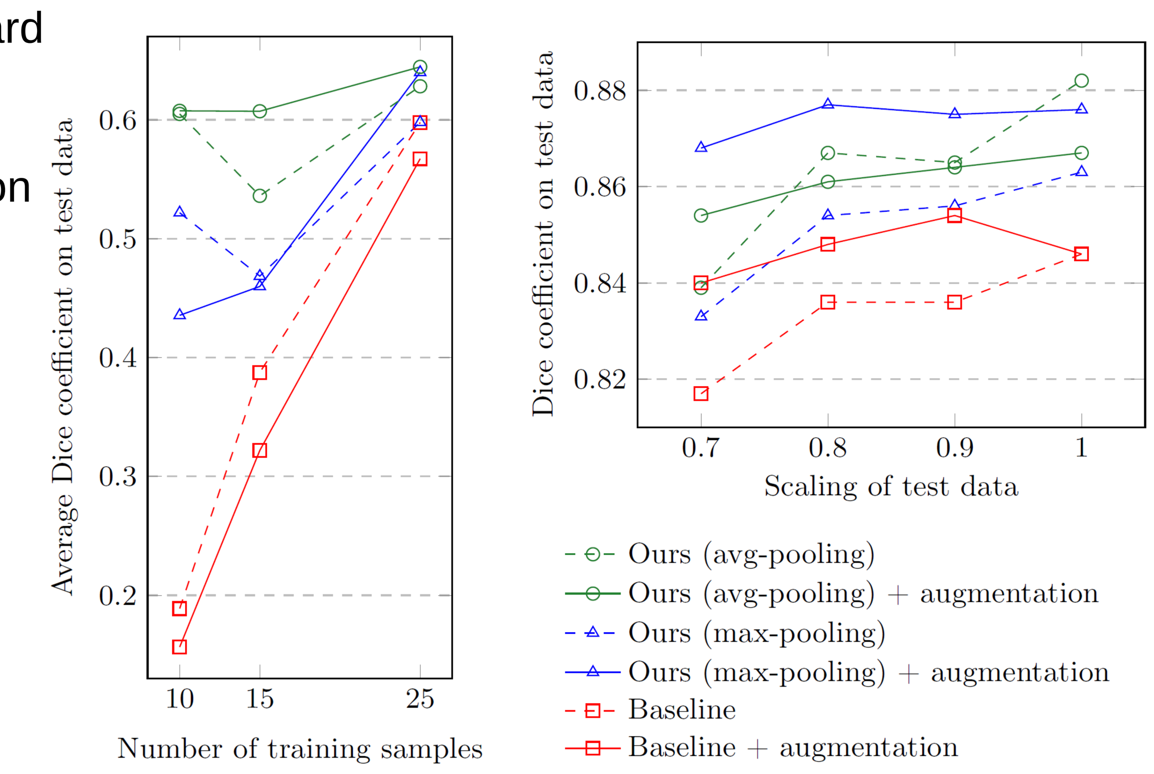
- Group-Convolution[1]: $[f \star_G \psi](g) = \int_{x \in X} f(x)\, L_g[\psi](x)\, d\mu(x)$
- Scale-equivariance has so far only been explored for 2D convolutions.
- We propose an extension of "Scale-Equivariant Steerable Networks"[2] to 3D.

Reference	Method	Optional Interscale Interaction	Admissible scales	Global Equivariance	Setting
Kanazawa, Sharma, and Jacobs 2014	Input scaling	No	Grid	No	2D
Xu, Xiao, Zhang, et al. 2014	Filter scaling	No	Grid	Yes	2D
Marcos, Kellenberger, Lobry, and Tuia 2018	Input scaling	Yes	Grid	Yes	2D
Worrall and Welling 2019	Filter Dilation	Yes	Integer	Yes	2D
Ghosh and Gupta 2019	Steerable Filters	No	All	No	2D
Sosnovik, Szmaja, and Smeulders 2019	Steerable Filters	Yes	All	Yes	2D
Zhu, Qiu, Calderbank, et al. 2019	Steerable Filters	Yes	All	Yes	2D
Bekkers 2019	Steerable Filters	Yes	All	Yes	2D
Naderi, Goli, and Kasaei 2020	Steerable Filters	Yes	All	Yes	2D
Sosnovik, Moskalev, and Smeulders 2021	Filter Dilation	Yes	$\sqrt{2}^i, i \in \mathbb{N}$.	Yes	2D
Lindeberg 2021	Steerable Filters	No	All	Yes	2D
Jansson and Lindeberg 2021	Steerable Filters	No	All	Yes	2D
Ours	**Steerable Filters**	**Yes**	**All**	**Yes**	**3D**

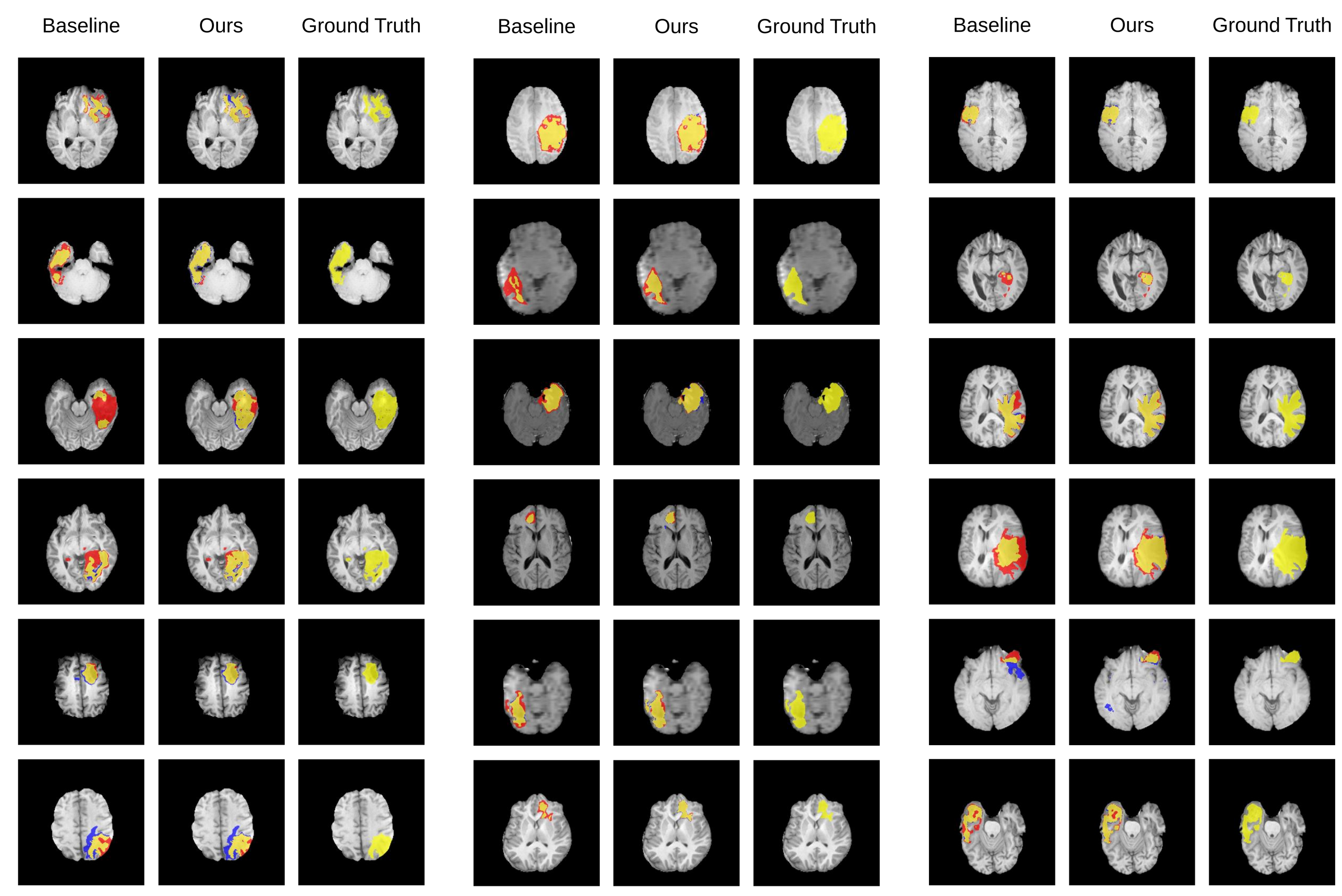
[1]Cohen, Taco, and Max Welling. "Group equivariant convolutional networks." *International conference on machine learning*. PMLR, 2016.
[2]Sosnovik, Ivan, Michał Szmaja, and Arnold Smeulders. "Scale-Equivariant Steerable Networks." *International Conference on Learning Representations*. 2019.

Experiments

- Comparison with standard convolutional baseline model on the MICCAI brain tumor segmentation dataset[1] (MRI scans).
- Evaluations with scaled training/test data and reduced training data sizes demonstrate the efficacy of our scale-equivariant layers.



Qualitative Results:



[1]Menze, Bjoern H., et al. "The multimodal brain tumor image segmentation benchmark (BRATS)." *IEEE transactions on medical imaging* 34.10 (2014): 1993-2024.

Conclusions

- We propose an extension of scale-equivariance to 3D convolutions.
- Our newly created network layers include scale-equivariant (transposed) convolutions, pooling and normalization layers.
- Our model demonstrates strong generalization in a low data setting and on scaled test data.

Future Work

- Extension to other 3D data representations (e.g., point clouds)
- Investigation of the feasibility of scale-equivariance for 6D dMRI and other imaging modalities