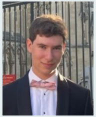
UNIVERSITÄT ZU LÜBECK
INSTITUTE OF MATHEMATICS AND IMAGE COMPUTING

CAMBRIDGE IMAGE ANALYSIS
DAMTP - UNIVERSITY OF CAMBRIDGE

Towards Diffeomorphism-Equivariant Neural Networks via Canonicalization (DiffeoNN)

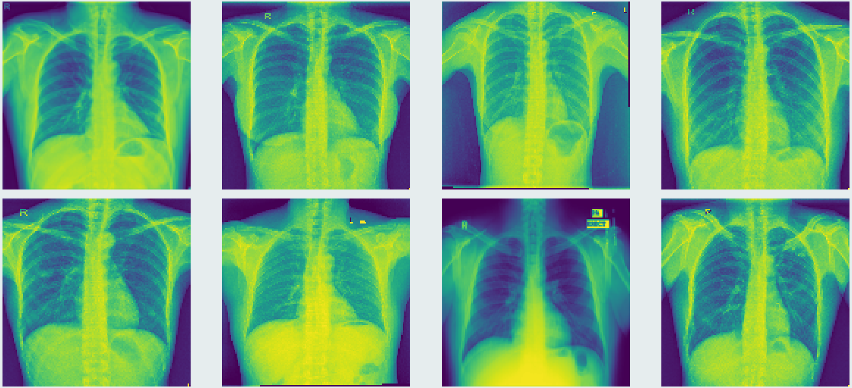
     

Josephine Elisabeth Oettinger[1,2], Zakhar Shumaylov[2], Peter Zaika[2]
Johannes Bostelmann[1], Jan Lellmann[1], Carola-Bibiane Schönlieb[2]

[1]Institute of Mathematics and Image Computing, University of Luebeck and [2]Department of Applied Mathematics and Theoretical Physics, University of Cambridge

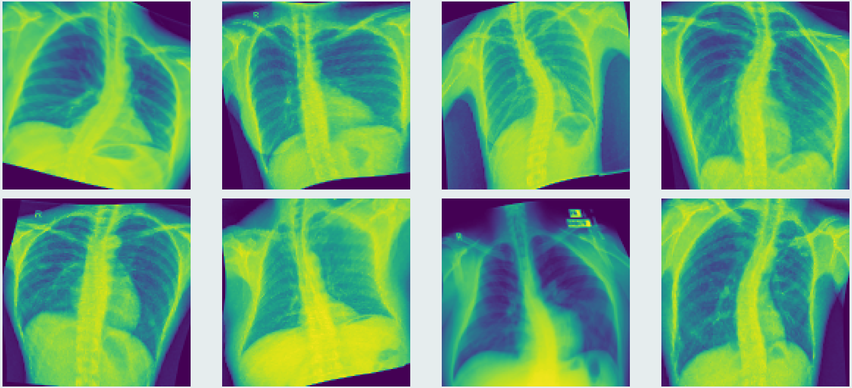
https://arxiv.org/abs/2602.06695

Standard CNNs are typically trained on well-aligned data...

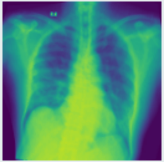


data: [Karmakar, 2024]

But, data naturally varies...

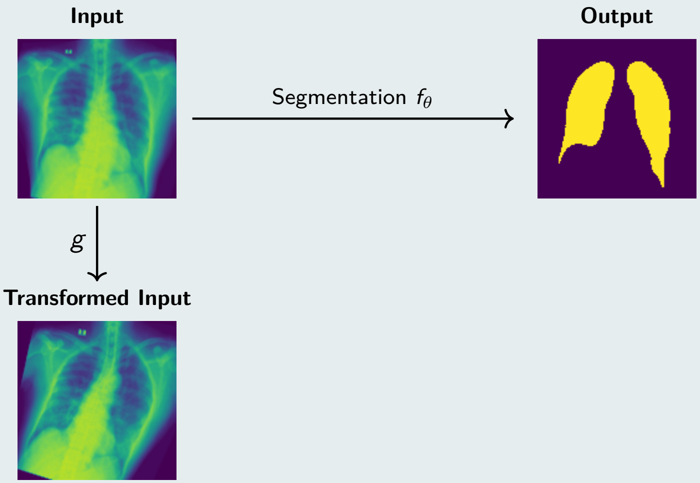


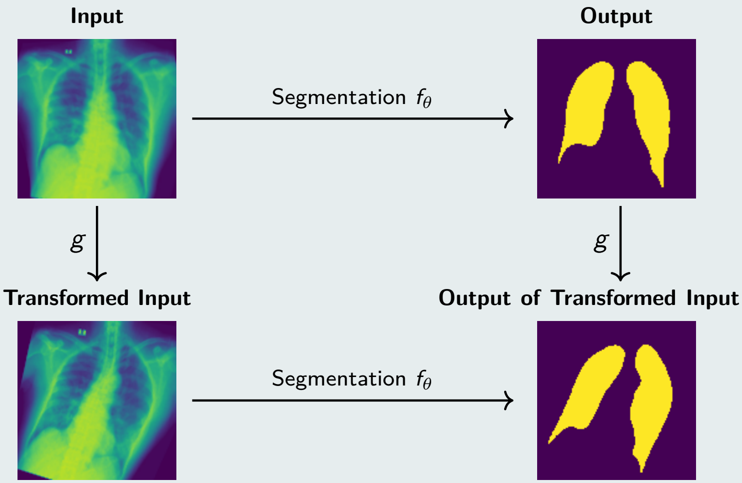
Input



Segmentation f_θ $\longrightarrow$

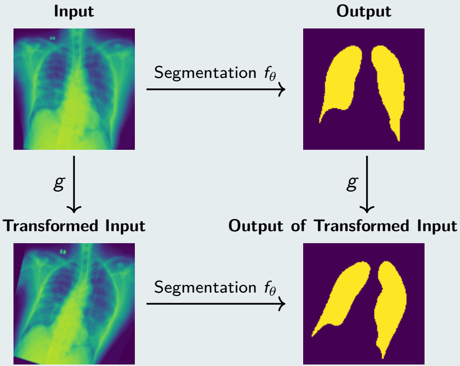
Output


Input
Segmentation f_θ
Output
g
Transformed Input


Input
Output
Segmentation f_θ
g
g
Transformed Input
Output of Transformed Input
Segmentation f_θ

Equivariance

G-Equivariant

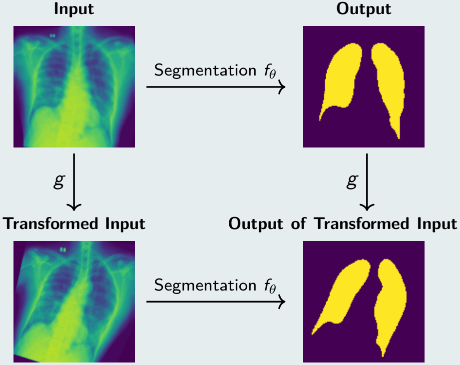


$$\tilde{f}_\theta(g \cdot x) = g \cdot \tilde{f}_\theta(x)$$

data: [RSUA, 2023]

Equivariance

G-Equivariant



$$\tilde{f}_\theta(g \cdot x) = g \cdot \tilde{f}_\theta(x)$$

Approaches

- Data augmentation [Shorten & Khoshgoftaar, 2019]
- Group-equivariant layers [Cohen & Welling, 2016]
- Averaging-based methods: group averaging [Murphy et al., 2019], frame averaging [Puny et al., 2022]
- **Canonicalization** [Kaba et al., 2023] [Shumaylov et al., 2024]

Diffeomorphisms

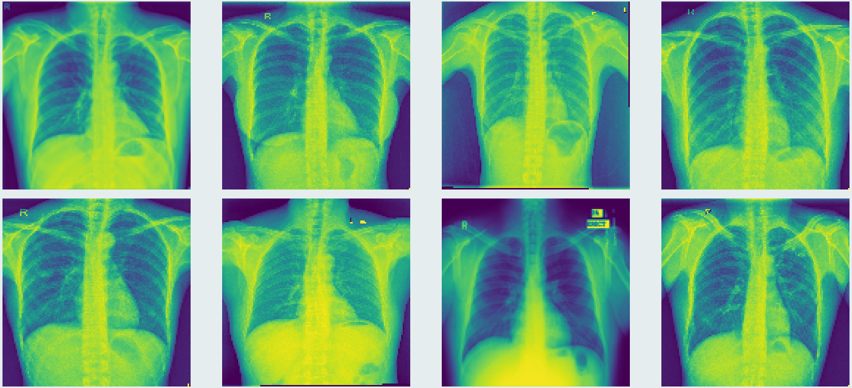
Definition (Diffeomorphism)

$\mathcal{X}$ and $\mathcal{Y}$ differentiable manifolds.
A map $g : \mathcal{X} \to \mathcal{Y}$ is called a diffeomorphism if

- g is bijective
- g is differentiable
- inverse $g^{-1} : \mathcal{Y} \to \mathcal{X}$ is differentiable

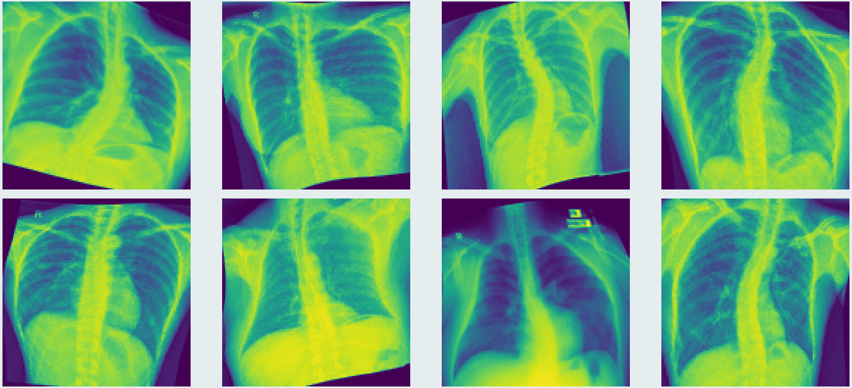
When $\mathcal{X} = \mathcal{Y}$, this set forms an infinite-dimensional group under composition, $\mathcal{D}(\mathcal{X})$.

Training Data X_E

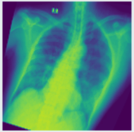


data: [Karmakar, 2024]

Input Images (Diffeomorphic Transformations of X_E)

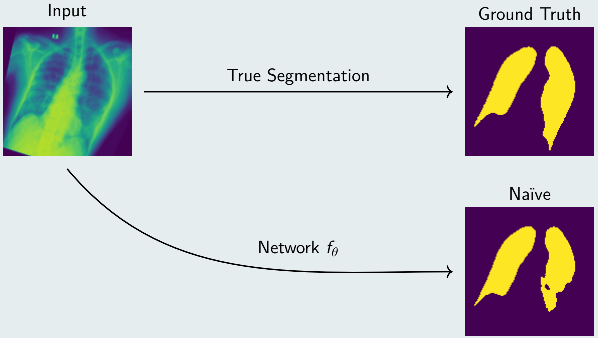


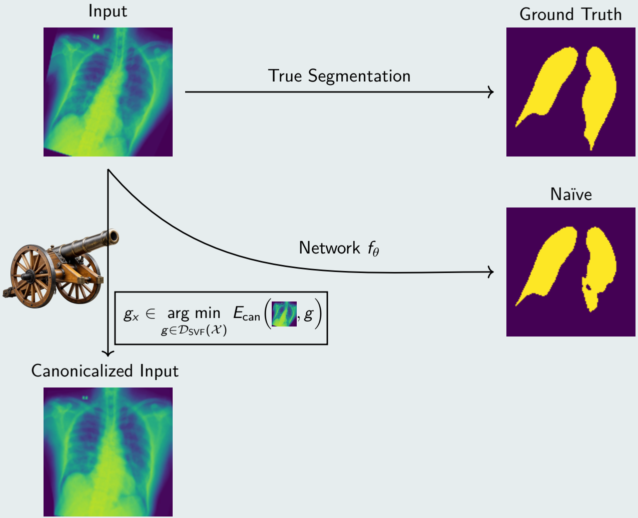
Input

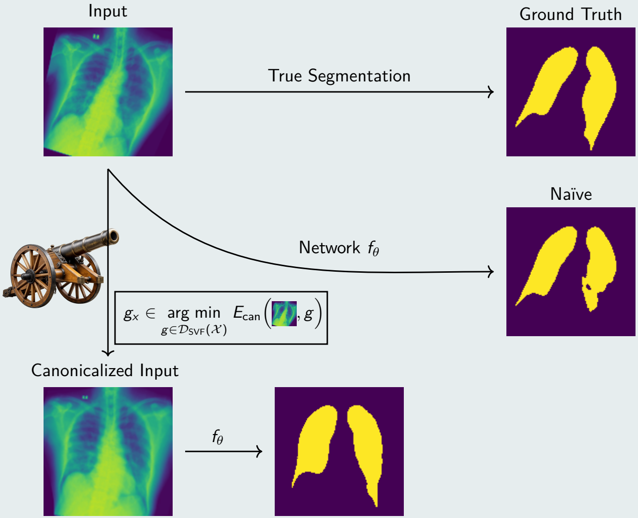


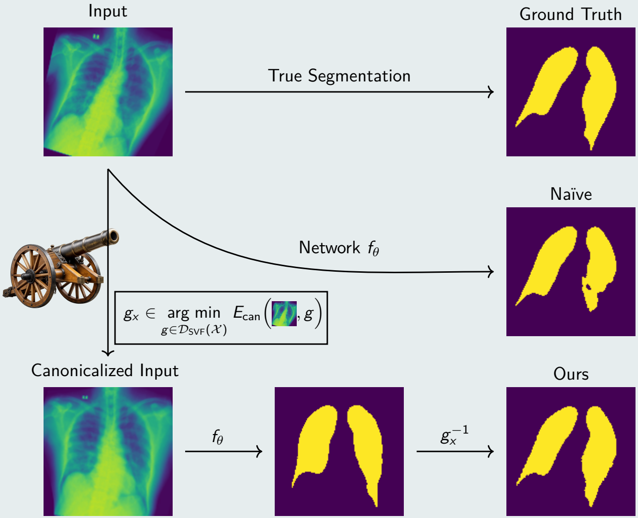
True Segmentation $\longrightarrow$

Ground Truth


Input
Ground Truth
True Segmentation
Naïve
Network f_θ

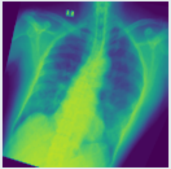

Input
Ground Truth
True Segmentation
Naïve
Network f_θ
$g_x \in \arg\min_{g \in \mathcal{D}_{\text{SVF}}(\mathcal{X})} E_{\text{can}}(\,\cdot\,, g)$
Canonicalized Input


Input
Ground Truth
True Segmentation
Naïve
Network f_θ
$g_x \in \arg\min_{g\in\mathcal{D}_{\mathrm{SVF}}(\mathcal{X})} E_{\mathrm{can}}(\cdot , g)$
Canonicalized Input
f_θ


Input
Ground Truth
True Segmentation
Naïve
Network f_θ
$g_x \in \arg\min_{g\in\mathcal{D}_{\mathrm{SVF}}(\mathcal{X})} E_{\mathrm{can}}(\,\cdot\,, g)$
Canonicalized Input
Ours
f_θ
g_x^{-1}

Canonicalization

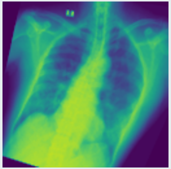
We have: input x.

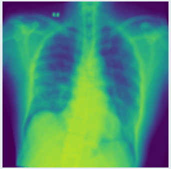


Canonicalization

We have: input x.

We want: canonical representation x_c, that is “close” to the training data X_E.





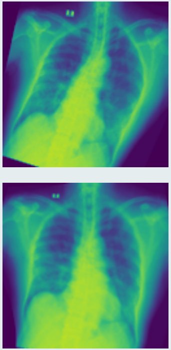
Canonicalization

We have: input x.

We want: canonical representation x_c, that is "close" to the training data X_E.

Idea: **energy-based canonicalization** [Shumaylov et al., 2024], canonicalize input to $x_c := g_x \cdot x$ with

$$g_x \in \operatorname*{arg\,min}_{g \in \mathcal{D}_{SVF}(\mathcal{X})} E_{\text{can}}(x, g).$$



Energy minimization through gradient-based optimization following [Bostelmann et al., 2024].

Canonicalization Energy E_{can}

Intuitively, energy $E_{\text{can}}(x, g)$ small when $g \cdot x$ is "close" to training data X_E and transformation g physically plausible, large otherwise.

Canonicalization Energy E_{can}

Intuitively, energy $E_{\mathrm{can}}(x, g)$ small when $g \cdot x$ is "close" to training data X_E and transformation g physically plausible, large otherwise.

$$E_{\mathrm{can}} : X \times \mathcal{D}_{\mathrm{SVF}}(X) \to \mathbb{R}, \qquad E_{\mathrm{can}}(x, g) := \underbrace{E_{X_E}(g \cdot x)}_{\text{image similarity}} + \underbrace{E_{\mathrm{reg}}(g)}_{\text{regularization}},$$

Canonicalization Energy E_{can}

Intuitively, energy $E_{\text{can}}(x, g)$ small when $g \cdot x$ is "close" to training data X_E and transformation g physically plausible, large otherwise.

$$E_{\text{can}} : X \times \mathcal{D}_{\text{SVF}}(X) \to \mathbb{R}, \qquad E_{\text{can}}(x, g) := \underbrace{E_{X_E}(g \cdot x)}_{\text{image similarity}} + \underbrace{E_{\text{reg}}(g)}_{\text{regularization}} ,$$

with

$$E_{X_E}(g \cdot x) := \lambda_{\text{VAE}} E_{\text{VAE}}(g \cdot x) + \lambda_{\text{adv}} E_{\text{adv}}(g \cdot x),$$

Canonicalization Energy E_{can}

Intuitively, energy $E_{\text{can}}(x, g)$ small when $g \cdot x$ is "close" to training data X_E and transformation g physically plausible, large otherwise.

$$E_{\text{can}} : X \times \mathcal{D}_{\text{SVF}}(X) \to \mathbb{R}, \qquad E_{\text{can}}(x, g) := \underbrace{E_{X_E}(g \cdot x)}_{\text{image similarity}} + \underbrace{E_{\text{reg}}(g)}_{\text{regularization}} ,$$

with

$$E_{X_E}(g \cdot x) := \lambda_{\text{VAE}} E_{\text{VAE}}(g \cdot x) + \lambda_{\text{adv}} E_{\text{adv}}(g \cdot x),$$

$$E_{\text{reg}}(g) := \underbrace{\lambda_{\text{grad}} \sum_{p \in \Omega} \sum_{i=1}^{d} (\nabla v_i(p))^2}_{\text{gradient loss}} + \underbrace{\lambda_{\text{jac}} \sum_{p \in \Omega} \max(0, -\det(\mathcal{J}_g(p)))}_{\text{Jacobian determinant loss}} .$$

SVF-Based Diffeomorphisms $\mathcal{D}_{\mathsf{SVF}}$

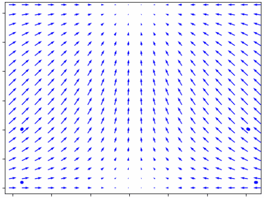
Stationary Velocity Field (SVF): vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: velocity of a fluid at position p.

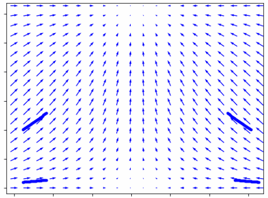
Flow: $\varphi_t(p) : \Omega \to \Omega$,

- defined as solution to ODE:

$$\frac{\mathrm{d}\varphi_t(p)}{\mathrm{d}t} = v(\varphi_t(p)), \quad \varphi_0(p) = p, \quad t \in [0, 1].$$



SVF-Based Diffeomorphisms $\mathcal{D}_{\mathsf{SVF}}$



Stationary Velocity Field (SVF): vector field $v : \Omega \to \mathbb{R}^d$,

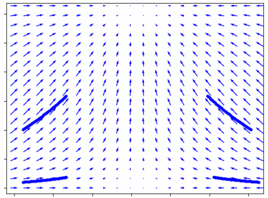
- $v(p)$: velocity of a fluid at position p.

Flow: $\varphi_t(p) : \Omega \to \Omega$,

- defined as solution to ODE:

$$\frac{\mathrm{d}\varphi_t(p)}{\mathrm{d}t} = v(\varphi_t(p)), \quad \varphi_0(p) = p, \quad t \in [0,1].$$

SVF-Based Diffeomorphisms $\mathcal{D}_{\text{SVF}}$



Stationary Velocity Field (SVF): vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: velocity of a fluid at position p.

Flow: $\varphi_t(p) : \Omega \to \Omega$,

- defined as solution to ODE:

$$\frac{\mathrm{d}\varphi_t(p)}{\mathrm{d}t} = v(\varphi_t(p)), \quad \varphi_0(p) = p, \quad t \in [0, 1].$$

SVF-Based Diffeomorphisms $\mathcal{D}_{\mathsf{SVF}}$

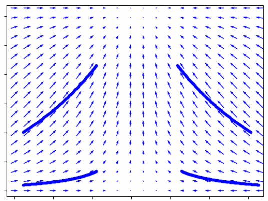
Stationary Velocity Field (SVF): vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: velocity of a fluid at position p.

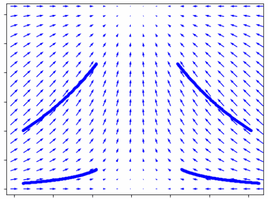
Flow: $\varphi_t(p) : \Omega \to \Omega$,

- defined as solution to ODE:

$$\frac{\mathrm{d}\varphi_t(p)}{\mathrm{d}t} = v(\varphi_t(p)), \quad \varphi_0(p) = p, \quad t \in [0, 1].$$



SVF-Based Diffeomorphisms $\mathcal{D}_{\text{SVF}}$



Stationary Velocity Field (SVF): vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: velocity of a fluid at position p.

Flow: $\varphi_t(p) : \Omega \to \Omega$,

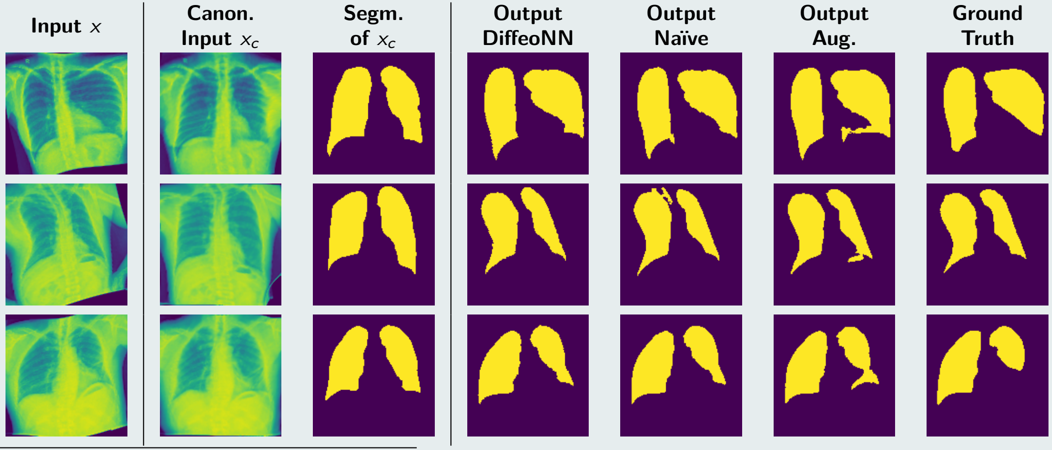
- defined as solution to ODE:

$$\frac{\mathrm{d}\varphi_t(p)}{\mathrm{d}t} = v(\varphi_t(p)), \quad \varphi_0(p) = p, \quad t \in [0, 1].$$

Diffeomorphism:

- solution of ODE at $t = 1$: $\varphi_1 = \exp(v)$,
- exponential map $\exp : \text{SVF} \to \mathcal{D}_{\text{SVF}}(X)$,
- structured and computationally efficient subset of full group of diffeomorphic transformations

Diffeomorphism-Equivariant Lung Segmentation

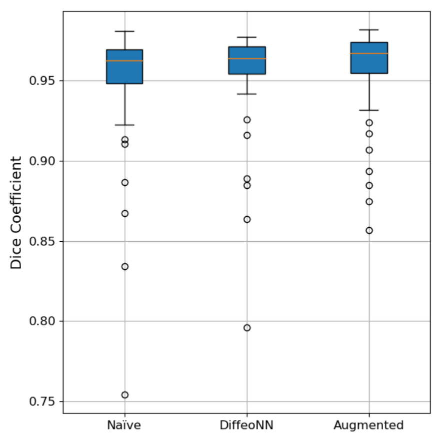


data: [Karmakar, 2024]

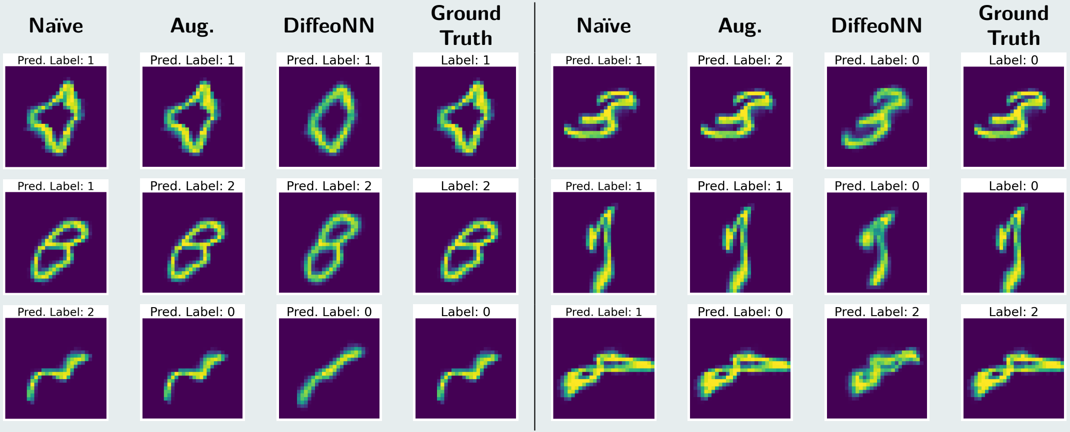
Diffeomorphism-Equivariant Lung Segmentation

Model	IoU ↑	Dice ↑	Acc. ↑
Naïve	0.9147	0.9547	0.9755
DiffeoNN (ours)	0.9214	0.9586	0.9777
Aug.	**0.9251**	**0.9606**	**0.9790**

Table: Mean Intersection-over-Union (IoU), Dice coefficient (Dice), and pixel-wise accuracy (Acc.) on diffeomorphically transformed **chest X-rays**.



Diffeomorphism-Invariant Homology Classification

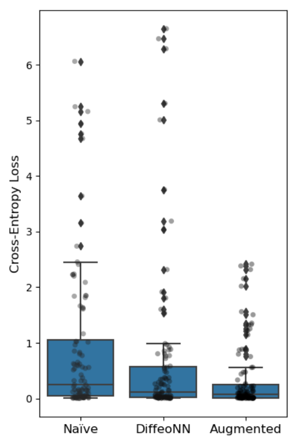


data: [LeCun et al., 1998]

Diffeomorphism-Invariant Homology Classification

MODEL	ACC. ↑
NAÏVE	0.68
DIFFEONN (OURS)	**0.82**
AUG.	**0.82**

Table: Mean accuracy (Acc.) on diffeomorphically transformed **MNIST**.



Bounding the Generalization Error

For energy-based canonicalization in LieLAC setup [Shumaylov et al., 2024]:
(compact and finite-dimensional setting):

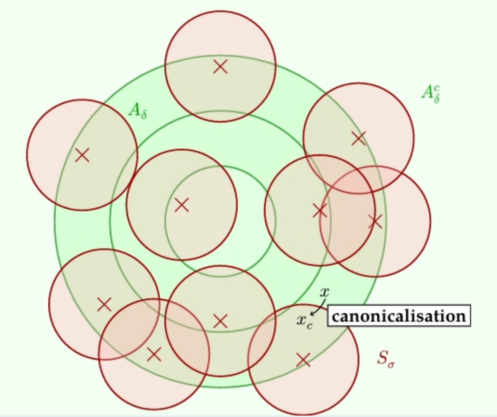
$$\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)), g_x \cdot y_x^{gt})] \leq \epsilon + (L_{\max} - \epsilon)\alpha,$$

with

- ϵ bounds training error,
- $L_{\max}$ maximal possible loss,
- α small, when X_E well sampled (probability mass on poorly-sampled group orbits),
- under mild assumption.

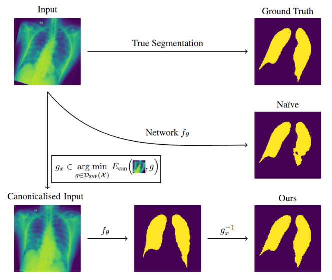
Proof Idea

$$
\begin{aligned}
&\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)),\ g_x \cdot y_x^{gt})] \\
&\quad = \int_{\mathcal{X}} L(f_\theta(g_x \cdot x), y^{gt}_{g_x \cdot x}) \mathrm{d}\mu_T(x) \\
&\quad = \underbrace{\int_{A_\delta} L(f_\theta(g_x \cdot x), y^{gt}_{g_x \cdot x}) \mathrm{d}\mu_T(x)}_{\leq (1-\mu_T(A_\delta^c))\epsilon} \\
&\qquad + \underbrace{\int_{A_\delta^c} L(f_\theta(g_x \cdot x), y^{gt}_{g_x \cdot x}) \mathrm{d}\mu_T(x)}_{\leq \mu_T(A_\delta^c) L_{\max}} \\
&\quad \leq \epsilon + (L_{\max} - \epsilon)\alpha
\end{aligned}
$$



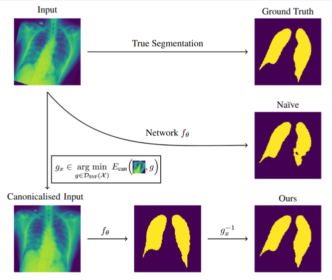
Bounding the Generalization Error

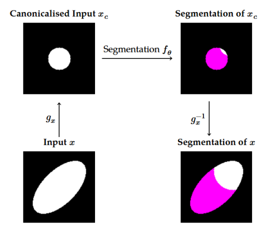
$$\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)), g_x \cdot y_x^{gt})] \\ \leq \epsilon + (L_{\max} - \epsilon)\alpha$$



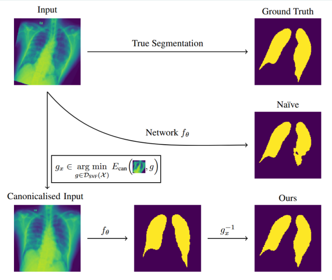
Bounding the Generalization Error

$$\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)), g_x \cdot y_x^{gt})] \\ \leq \epsilon + (L_{\max} - \epsilon)\alpha$$

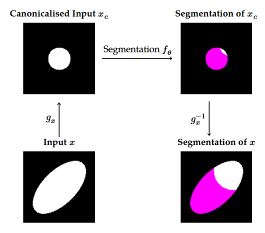




Bounding the Generalization Error

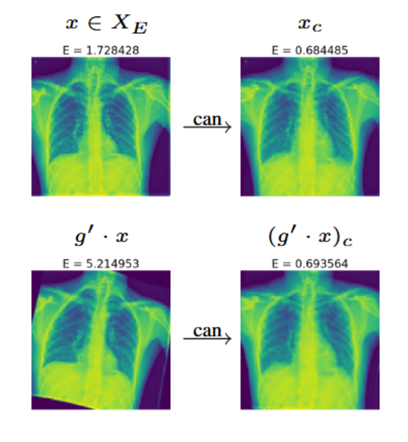


$$\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)),\, g_x \cdot y_x^{gt})] \\ \leq \epsilon + (L_{\max} - \epsilon)\alpha$$



$$\mathbb{E}_{\mu_T}[L(g_x^{-1} \cdot (f_\theta(g_x \cdot x)),\; y_x^{gt})] \\ \leq C \cdot \mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)),\; g_x \cdot y_x^{gt})].$$

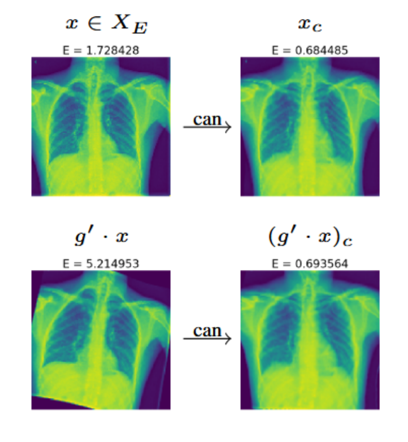
Invariance of Canonicalization and Empirical Equivariance Error



If $\tilde{f}_\theta$ does make the inner network f_θ diffeomorphism-equivariant, for all images from the training dataset $x \in X_E$:

$$f_\theta(x) = \tilde{f}_\theta(x) = g_x^{-1} \cdot f_\theta(g_x \cdot x)$$

Invariance of Canonicalization and Empirical Equivariance Error

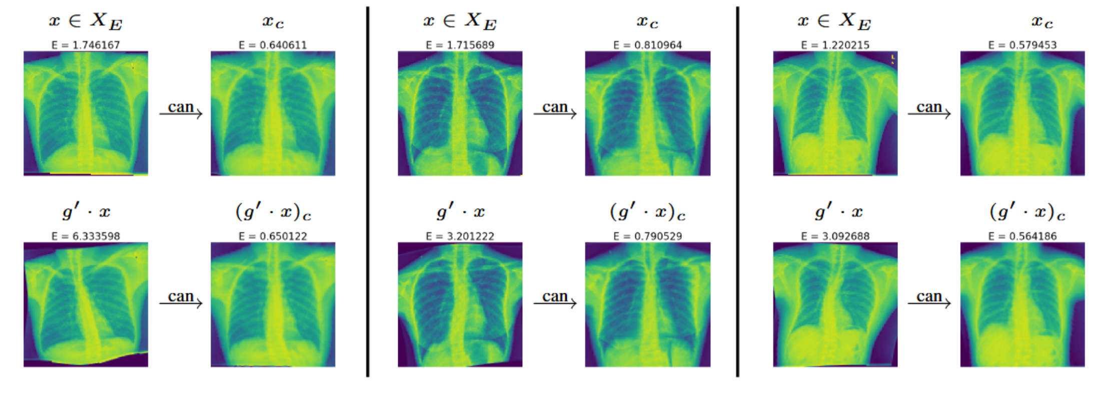


If $\tilde{f}_\theta$ does make the inner network f_θ diffeomorphism-equivariant, for all images from the training dataset $x \in X_E$:

$$f_\theta(x) = \tilde{f}_\theta(x) = g_x^{-1} \cdot f_\theta(g_x \cdot x)$$

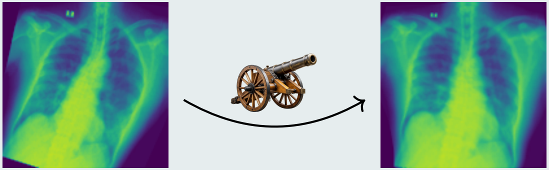
- mean Dice coefficient
  $\mathrm{Dice}(f_\theta(x), \tilde{f}_\theta(x)) = 0.9812$
- mean pixel-wise accuracy
  $\mathrm{Acc}(f_\theta(x), \tilde{f}_\theta(x)) = 0.9902$

More Examples on Invariance of Canonicalization



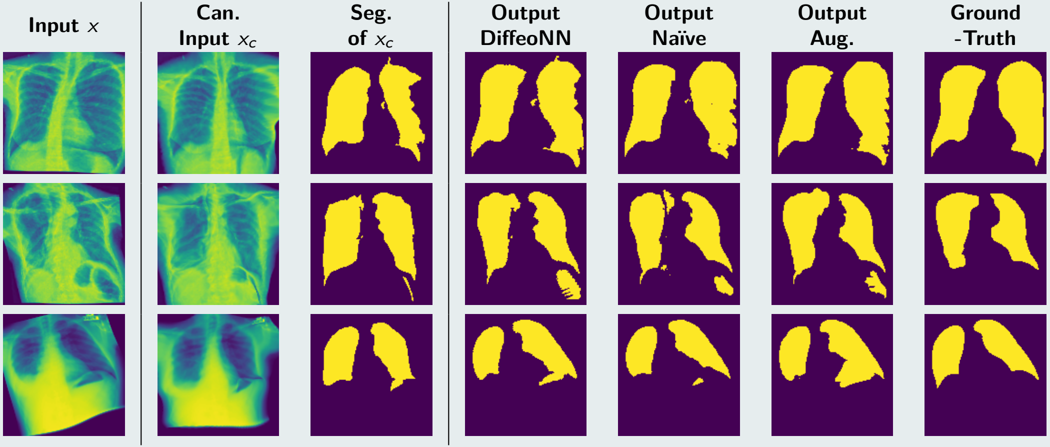
Diffeomorphism-Equivariant Neural Networks (DiffeoNN)

- turn any network diffeomorphism-equivariant
- no data augmentation or retraining necessary (only training on original training dataset X_E)
- bounded generalization error

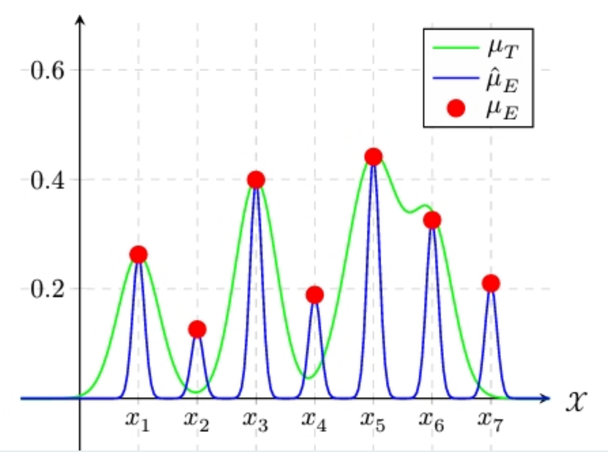


additional slides...

Lung Segmentation: Failed Canonicalization



Bounding Generalization Error - Measures



References

- [Shumaylov et al., 2024] Zakhar Shumaylov, Peter Zaika, James Rowbottom, Ferdia Sherry, Melanie Weber, and Carola-Bibiane Schönlieb. "Lie algebra canonicalization: Equivariant neural operators under arbitrary lie groups.", 2024. arXiv preprint arXiv:2410.02698.
- [Bostelmann et al., 2024] Johannes Bostelmann, Ole Gildemeister, and Jan Lellmann. "Stationary velocity fields on matrix groups for deformable image registration.", 2024. arXiv preprint arXiv:2410.10997.
- [LeCun et al., 1998] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. "Gradient-based learning applied to document recognition." Proc. IEEE, 86:2278–2324, 1998. URL https://api.semanticscholar.org/CorpusID:14542261.
- [RSUA, 2023] RSUA, R. RSUA Chest X-Ray Dataset. Version V1. 2023. doi: 10.17632/2jg8vfdmpm. url: https://doi.org/10.17632/2jg8vfdmpm.1.
- [Karmakar, 2024] Karmakar, Tapendu. Chest X-ray Dataset for Lung Segmentation. Kaggle, 2024, https://www.kaggle.com/datasets/iamtapendu/chest-x-ray-lungs-segmentation. Accessed 13 Feb. 2026.