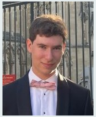
UNIVERSITÄT ZU LÜBECK
INSTITUTE OF MATHEMATICS AND
IMAGE COMPUTING

CAMBRIDGE IMAGE ANALYSIS
DAMTP - UNIVERSITY OF CAMBRIDGE

# Diffeomorphism-Equivariant Neural Networks

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

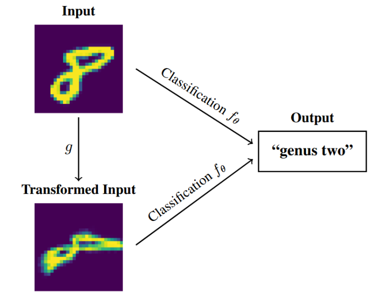
[1]Institute of Mathematics and Image Computing, University of Lübeck and [2]Department of Applied Mathematics and Theoretical Physics, University of Cambridge

17. March 2026

GAMM: Mathematical signal and image processing

# Invariance and Equivariance

## G-Invariant



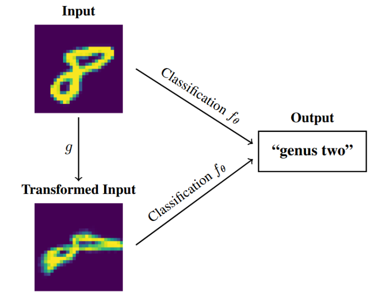

$$\tilde{f}_\theta(g \cdot x) = \tilde{f}_\theta(x),$$

## G-Equivariant

data/images: [LeCun et al., 1998], [RSUA, 2023]

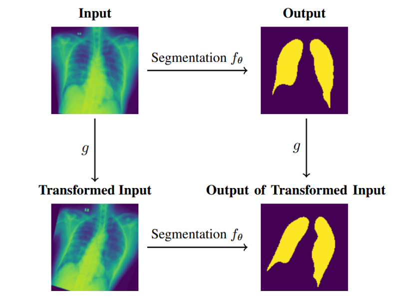
# Invariance and Equivariance

## G-Invariant




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

## G-Equivariant




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

data/images: [LeCun et al., 1998], [RSUA, 2023]

# Diffeomorphisms

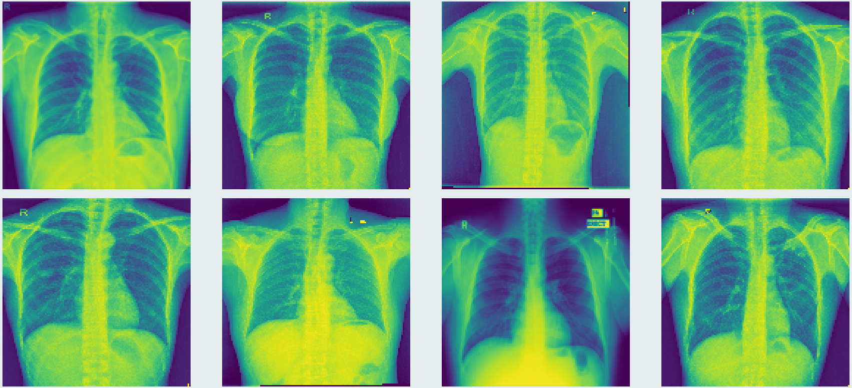
## Definition (Diffeomorphism)

$\mathcal{X}$ 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

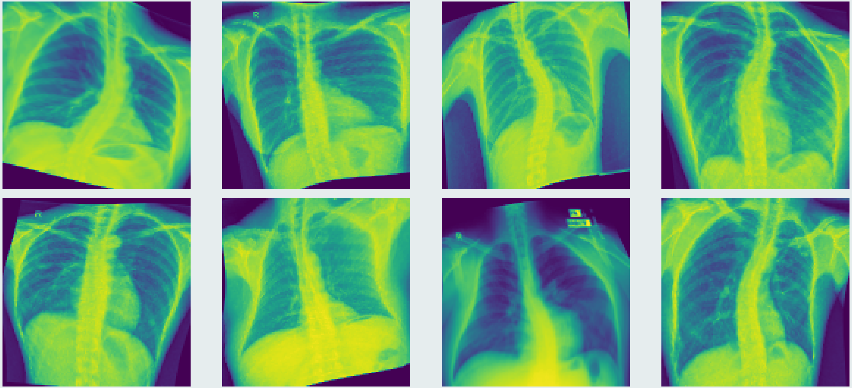
When $\mathcal{X} = \mathcal{Y}$, this set forms an infinite-dimensional group under composition, $\mathcal{D}(\mathcal{X})$.

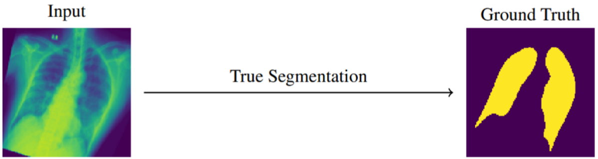
# Training Data $X_E$

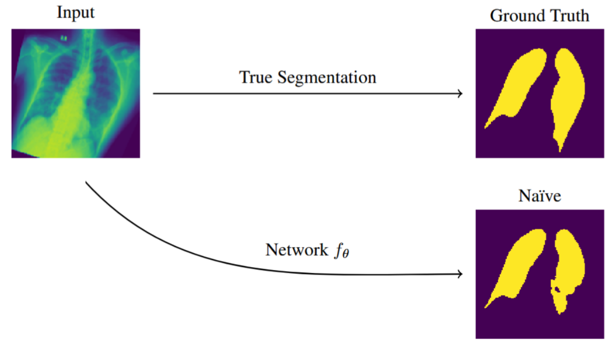


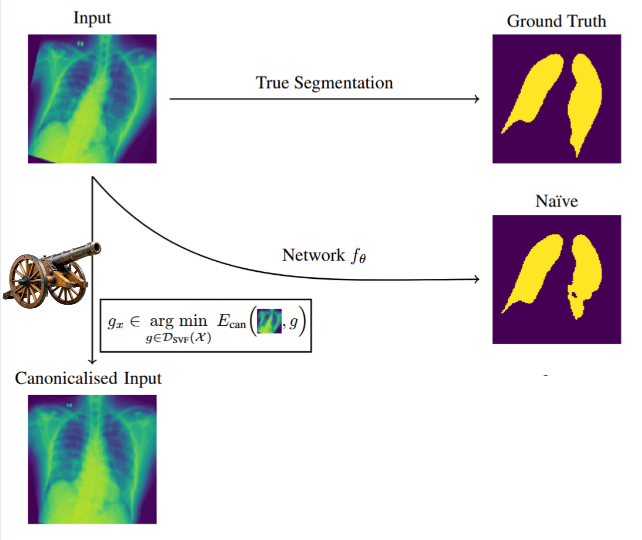
data: [Karmakar, 2024]

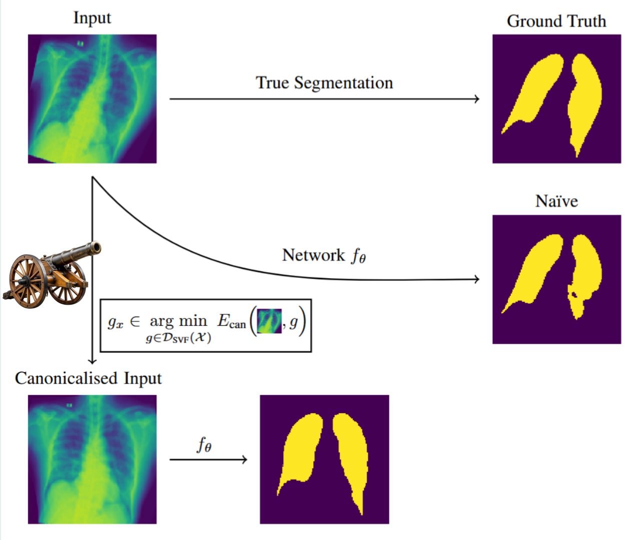
# Diffeomorphically Transformed Images

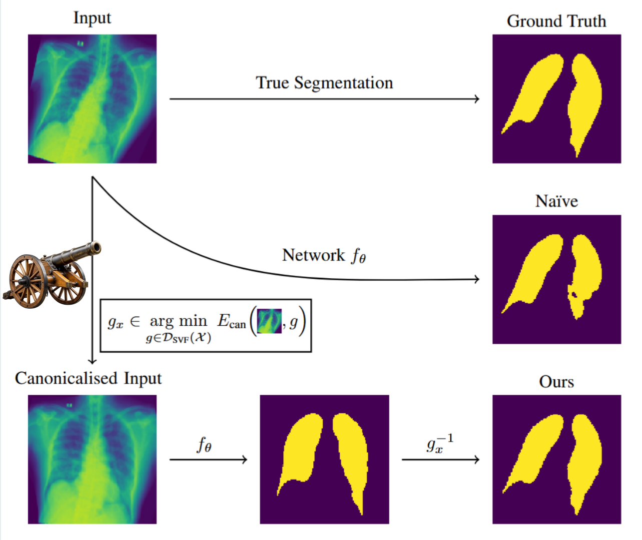



Input
True Segmentation
Ground Truth


Input
Ground Truth
True Segmentation
Naïve
Network $f_\theta$

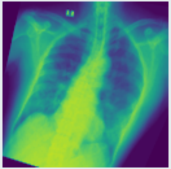

Input
Ground Truth
True Segmentation
Naïve
Network $f_\theta$
$g_x \in \arg\min_{g \in \mathcal{D}_{\mathrm{SVF}}(\mathcal{X})} E_{\mathrm{can}}(\cdot, g)$
Canonicalised Input


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


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

# Canonicalisation

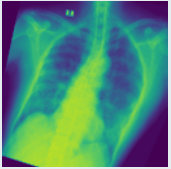
**We have:** input $x$.

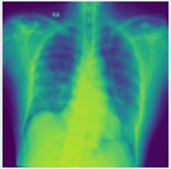


# Canonicalisation

**We have:** input $x$.

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





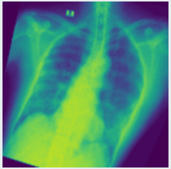
# Canonicalisation

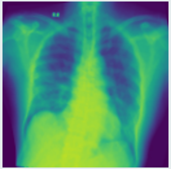
**We have:** input $x$.

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

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

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





Energy minimisation through gradient-based optimisation following [Bostelmann et al., 2024].

# Canonicalisation Energy $E_{\text{can}}$

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.

# Canonicalisation Energy $E_{\text{can}}$

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.

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

# Canonicalisation Energy $E_{\text{can}}$

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.

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

with

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

# Canonicalisation Energy $E_{\text{can}}$

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.

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

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 $\mathcal{D}_{\text{SVF}}$

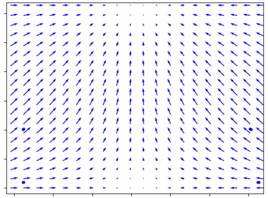
**Stationary Velocity Field (SVF)**: vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: 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 $\mathcal{D}_{\text{SVF}}$

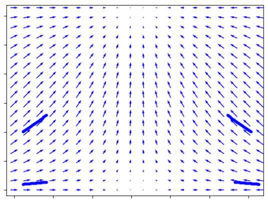
**Stationary Velocity Field (SVF)**: vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: 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 $\mathcal{D}_{\mathsf{SVF}}$

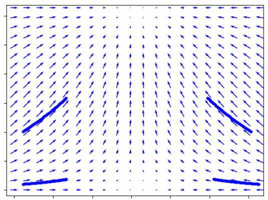
**Stationary Velocity Field (SVF)**: vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: 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 $\mathcal{D}_{\mathsf{SVF}}$

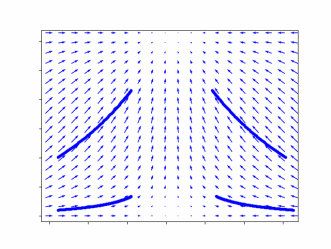
**Stationary Velocity Field (SVF)**: vector field $v : \Omega \to \mathbb{R}^d$,

- $v(p)$: 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 $\mathcal{D}_{\mathsf{SVF}}$

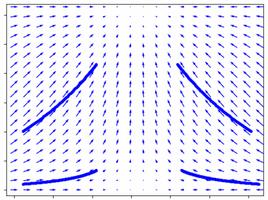
**Stationary Velocity Field (SVF)**: vector field $v : \Omega \to \mathbb{R}^d$,

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

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

- defined as solution to ODE:

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



**Diffeomorphism:**

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

# Diffeomorphism-Equivariant Lung Segmentation

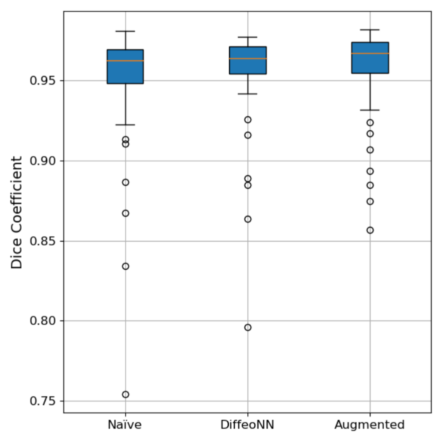
| Input $x$ | Canon. Input $x_c$ | Segm. of $x_c$ | Output DiffeoNN | Output Naïve | Output Aug. | Ground Truth |
|---|---|---|---|---|---|---|

data: [Karmakar, 2024]

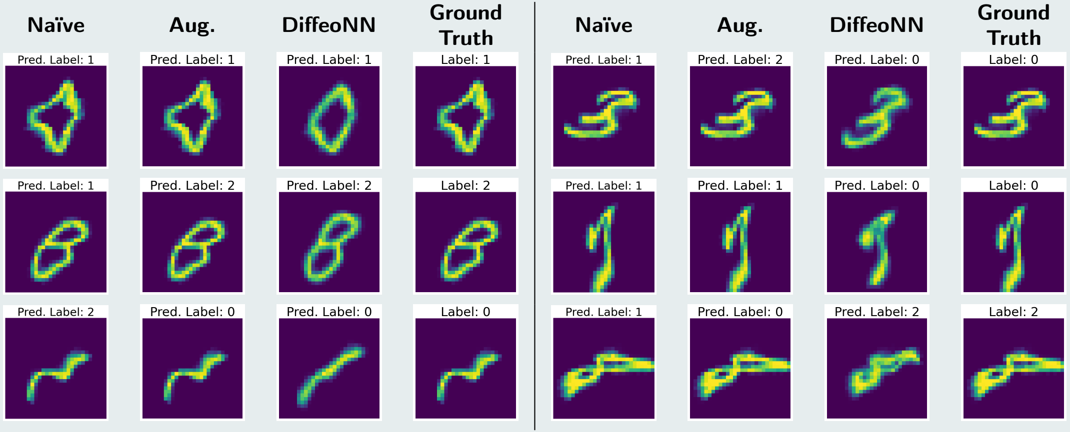
# Diffeomorphism-Equivariant Lung Segmentation

| 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

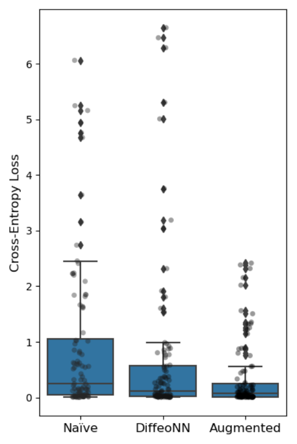



data: [LeCun et al., 1998]

# 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 Generalisation Error

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

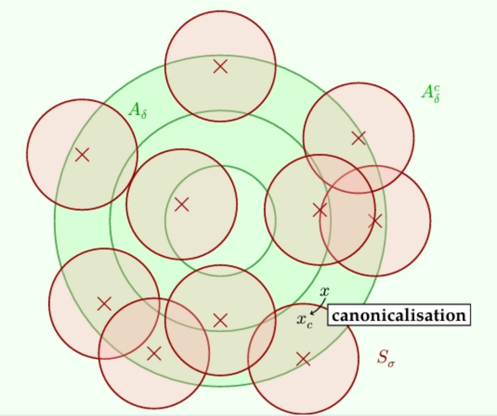
$$\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)), g_x \cdot y_x^{gt})] \leq \epsilon + (L_{\max} - \epsilon)\alpha,$$

with

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

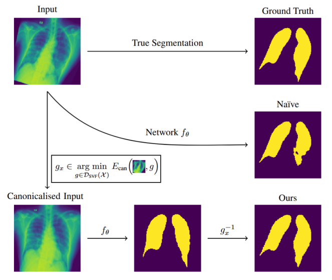
# Proof Idea

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



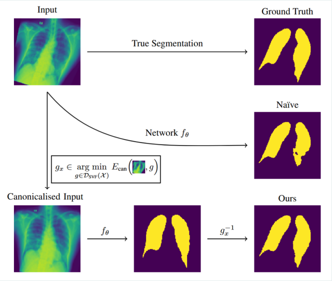
# Bounding the Generalisation Error

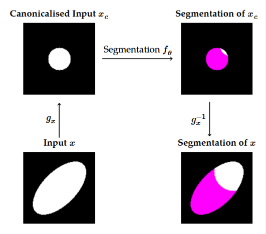
$$\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)), g_x \cdot y_x^{gt})] \\ \leq \epsilon + (L_{\max} - \epsilon)\alpha$$



# Bounding the Generalisation Error

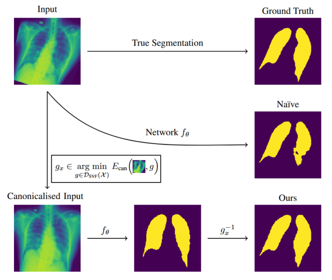
$$\mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x)), g_x \cdot y_x^{gt})] \leq \epsilon + (L_{\max} - \epsilon)\alpha$$

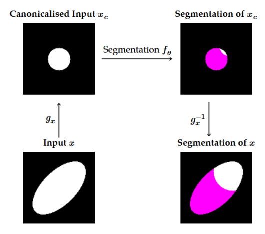


# Bounding the Generalisation Error

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

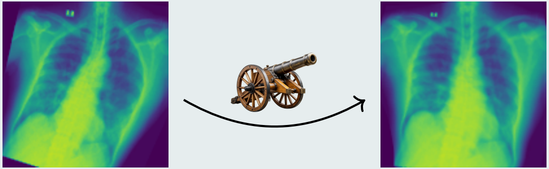


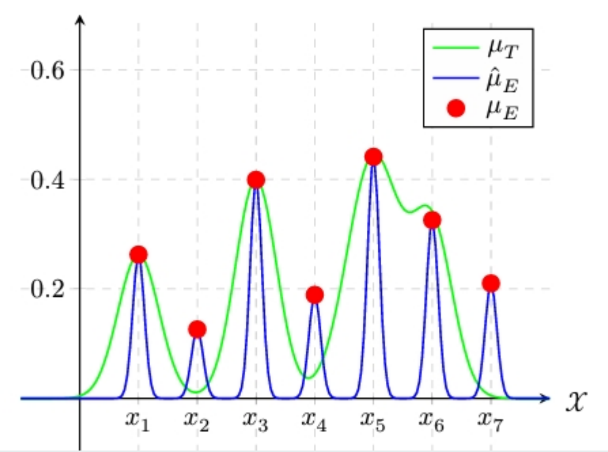

$$\mathbb{E}_{\mu_T}[L(g_x^{-1} \cdot (f_\theta(g_x \cdot x)), \ y_x^{gt})] \leq C \cdot \mathbb{E}_{\mu_T}[L(f_\theta(g_x \cdot x), \ g_x \cdot y_x^{gt})].$$

# Diffeomorphism-Equivariant Neural Networks (DiffeoNN)

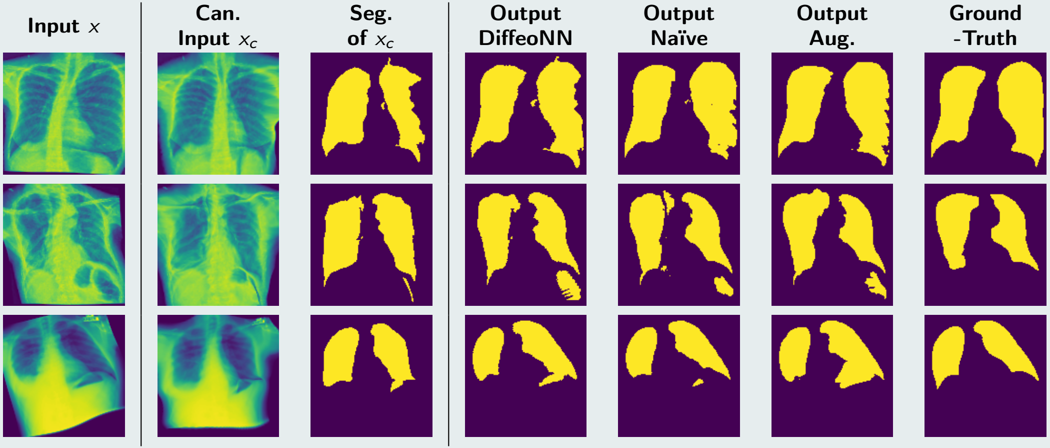
- turn any network diffeomorphism-equivariant
- no data augmentation or retraining necessary (only training on $X_E$)
- bounded generalisation error



# Bounding Generalisation Error - Measures



# Lung Segmentation: Failed Canonicalisation



# 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.