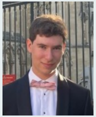
UNIVERSITÄT ZU LÜBECK
INSTITUTE OF MATHEMATICS AND IMAGE COMPUTING

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]

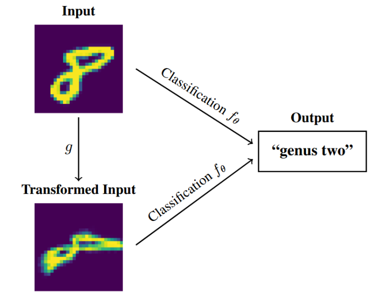
[1]Institute of Mathematics and Image Computing, University of Lübeck and [2]Department of Applied Mathematics and Theoretical Physics, University of Cambridge

17. March 2026

GAMM: Mathematical signal and image processing

Invariance and Equivariance

G-Invariant



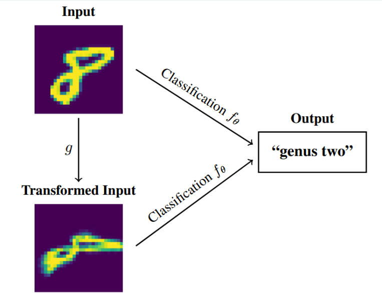
$$f(g \quad x) = f(x);$$

G-Equivariant

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

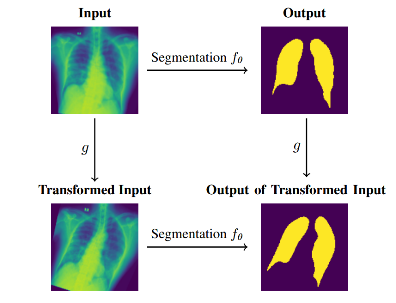
Invariance and Equivariance

G-Invariant



$$f(g \ \ x) = f(x);$$

G-Equivariant



$$f(g \ \ x) = g \ \ f(x);$$

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

Diffeomorphisms

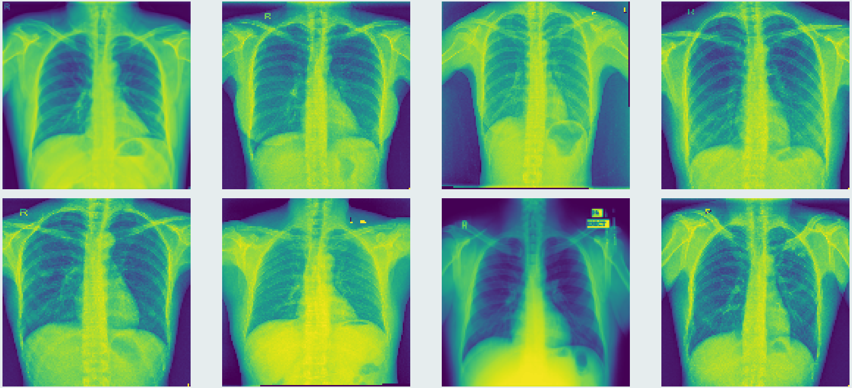
Definition (Diffeomorphism)

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

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

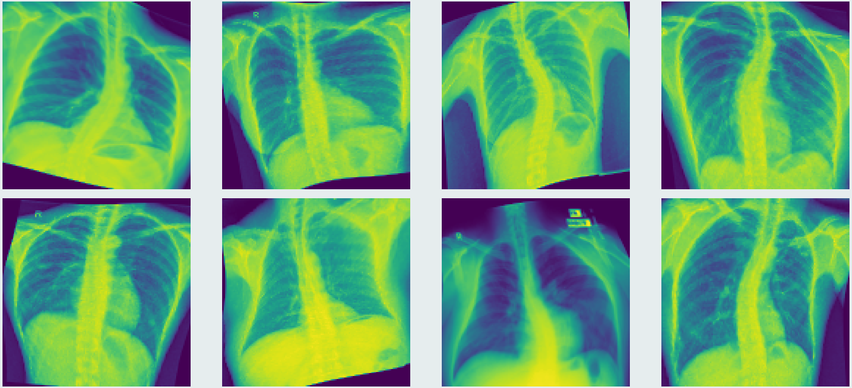
When $X = Y$, this set forms an infinite-dimensional group under composition, $D(X)$.

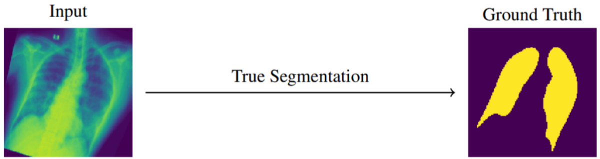
Training Data X_E

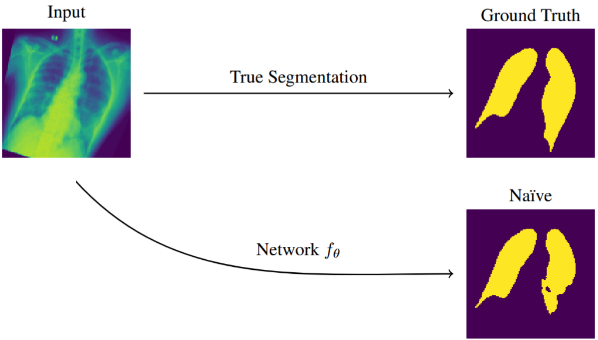


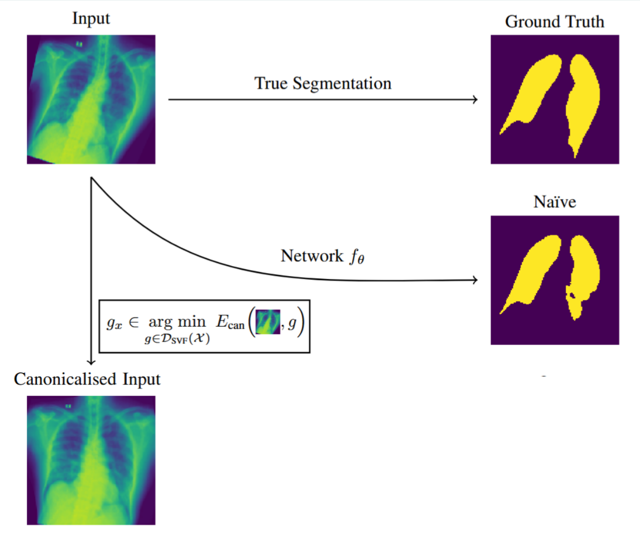
data: [Karmakar, 2024]

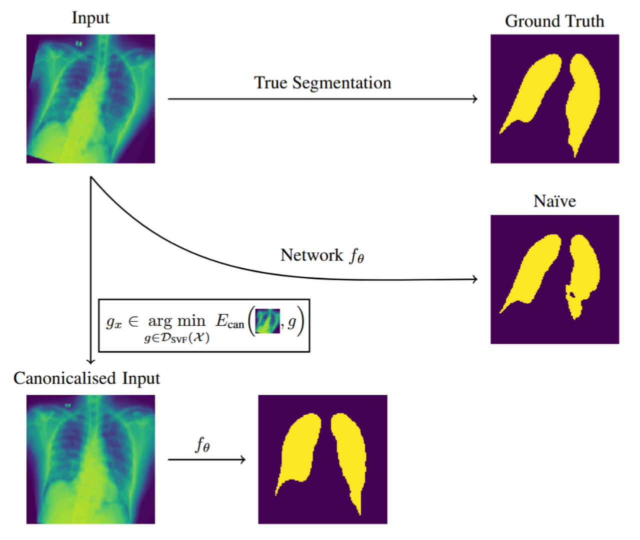
Diffeomorphically Transformed Images

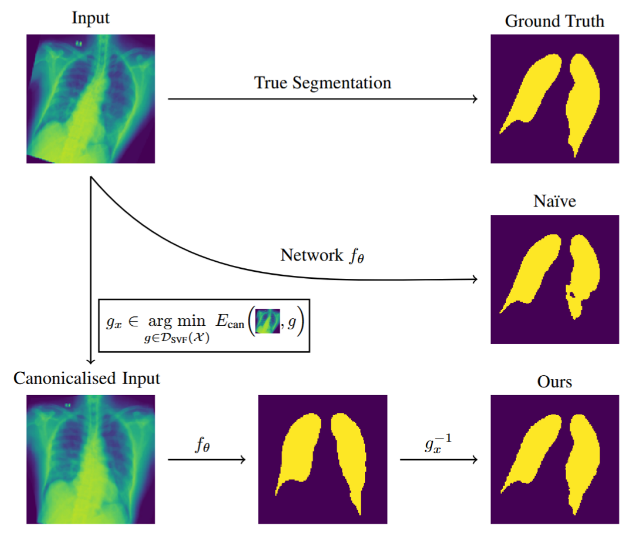



Input
True Segmentation
Ground Truth


Input
Ground Truth
True Segmentation
Naïve
Network f_θ


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


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


Input
Ground Truth
True Segmentation
Naïve
Network f_θ
$g_x \in \arg\min_{g \in \mathcal{D}_{SVF}(\mathcal{X})} E_{can}(\cdot , g)$
Canonicalised Input
Ours
f_θ
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 $x_c := g_x \; x$ with

$$g_x \in \arg\min_{g \in D_{SVF}(X)} E_{\mathrm{can}}(x; g).$$

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

Canonicalisation Energy E_{can}

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

Canonicalisation Energy E_{can}

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

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

Canonicalisation Energy E_{can}

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

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

with

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

Canonicalisation Energy E_{can}

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

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

with

$$E_{X_E}(g \circ x) := \lambda_{VAE} E_{VAE}(g \circ x) + \lambda_{adv} E_{adv}(g \circ x);$$

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

SVF-Based Diffeomorphisms D_{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:

$$\frac{d\varphi_t(p)}{dt} = v(\varphi_t(p)); \quad \varphi_0(p) = p; \quad \forall t \in [0;1]:$$

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

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:

$$\frac{d\varphi_t(p)}{dt} = v(\varphi_t(p)); \quad \varphi_0(p) = p; \quad t \in [0; 1]:$$

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

- 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:

$$\frac{d\varphi_t(p)}{dt} = v(\varphi_t(p)); \quad \varphi_0(p) = p; \quad t \in [0; 1]:$$

SVF-Based Diffeomorphisms D_{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:

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

SVF-Based Di eomorphisms D_{SVF}

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

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

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

- de ned as solution to ODE:

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

Di eomorphism:

- solution of ODE at $t = 1$: $\varphi_1 = \exp(v)$,
- exponential map $\exp : SVF \to D_{SVF}(X)$,
- structured and computationally e cient subset of full group of di eomorphic transformations

Di eomorphism-Equivariant Lung Segmentation

Input x	Canon. Input x_c	Segm. of x_c	Output Di eoNN	Output Naïve	Output Aug.	Ground Truth

data: [Karmakar, 2024]

Di eomorphism-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 coe cient (Dice), and pixel-wise accuracy (Acc.) on di eomorphically transformed chest X-rays.

Di eomorphism-Invariant Homology Classi cation

Naïve	Aug.	Di eoNN	Ground Truth	Naïve	Aug.	Di eoNN	Ground Truth

data: [LeCun et al., 1998]

Di eomorphism-Invariant Homology Classi cation

Model	Acc. "
Naïve	0.68
DiffeoNN (ours)	0.82
Aug.	0.82

Table: Mean accuracy (Acc.) on di eomorphically transformed MNIST.

Bounding the Generalisation Error

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

$$E_{\ T}[L(f\ (g_x\ \ x)); g_x\ \ y_x^{gt})]\quad + (L\ \ _{max}\ \ )\ ;$$

with

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

Proof Idea

$$E_{\;T}[L(f(g_x \;\; x)); \; g_x \;\; y_x^{gt})]$$

$$= \int_X L(f \; (g_x \;\; x); y^{gt}_{g_x \; x})d \;_T(x)$$

$$= \int_A \underbrace{L(f \; (g_x \;\; x); y^{gt}_{g_x \; x})d \;_T(x)}_{(1 \qquad _T(A^c))}$$

$$+ \int_{A^c} \underbrace{L(f \; (g_x \;\; x); y^{gt}_{g_x \; x})d \;_T(x)}_{_T(A^c)L_{max}}$$

$$+ (L \quad _{max} \quad)$$

Bounding the Generalisation Error

$$E_{T}[L(f\ (g_x\ \ x)); g_x\ \ y_x^{gt})]$$
$$+ (L\ \ _{max}\ \ )$$

Bounding the Generalisation Error

$$E \;_{T}[L(f\;(g_x\;\;x));\, g_x\;\;y_x^{gt})]$$

$$+\,(L\;\;_{max}\;\;)$$

Bounding the Generalisation Error

$$E_{T}[L(f\ (g_x\ \ x));\ g_x\ \ y_x^{gt})]$$
$$+\ (L\ \ _{max}\ \ )$$

$$E_{T}[L(g_x^{1}\ \ (f\ (g_x\ \ x));\ y_x^{gt})]$$
$$C\ \ E\ \ _{T}[L(f\ (g_x\ \ x);\ g_x\ \ y_x^{gt})]:$$

Di eomorphism-Equivariant Neural Networks (Di eoNN)

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

Bounding Generalisation Error - Measures

Lung Segmentation: Failed Canonicalisation

Input x	Can. Input x_c	Seg. of x_c	Output Di eoNN	Output Naïve	Output Aug.	Ground -Truth

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.