Analysis of (sub)Riemannian PDE-G-CNNs with New Kernel Approximations

Gijs Bellaard, Supervisor: Remco Duits

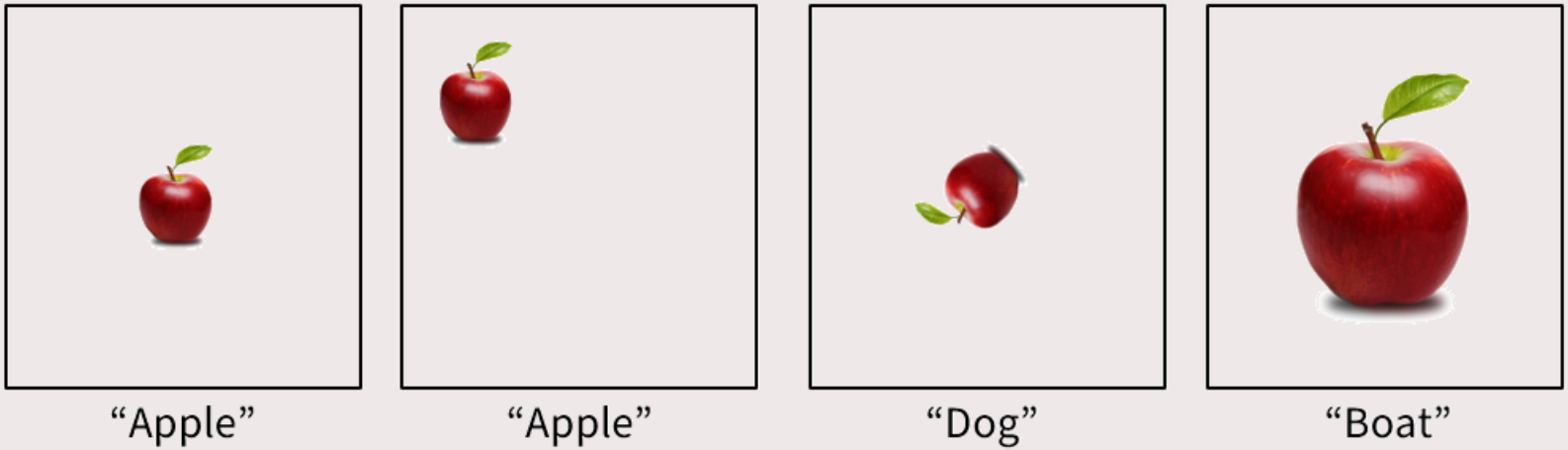
Dep. of Mathematics and Computer Science

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

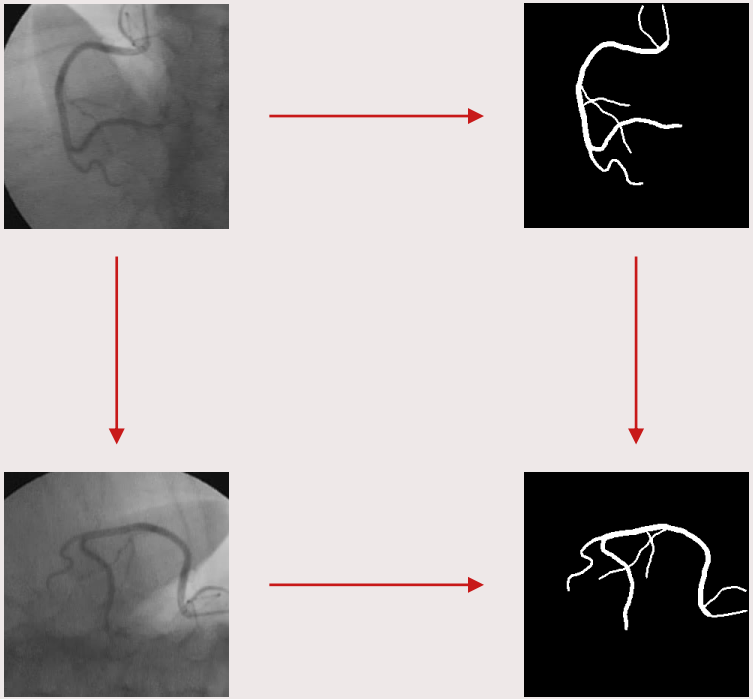
- G-CNNs
- The Homogeneous Space of Positions and Orientations
- PDE-G-CNNs
- The Logarithmic Approximation
- How good are the Approximations?

G-CNNs

Invariant Classification Network



Equivariant Segmentation Network



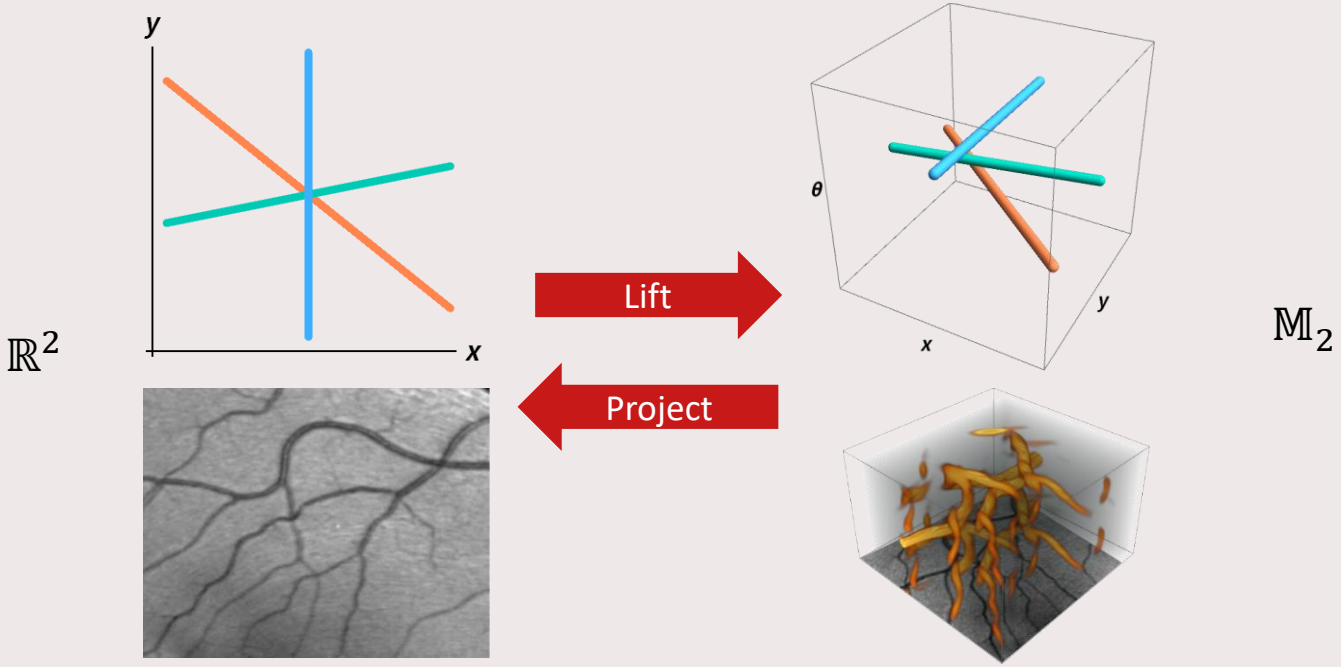
Group Equivariant Convolutional Neural Networks

- Introduced by Cohen and Welling in their homonymous paper from 2016.
- They create a classification network invariant under translations, mirror reflections, and quarter rotations.
- Their network ".. almost halves the error rate of the previous state of the art (2.28% vs 3.98% error)" when used on the RotNIST dataset.
- However, the data needed to be lifted to the higher dimensional space of positions and orientations.

Lifting

Image

Orientation Score



$\mathbb{R}^2$

$\mathbb{M}_2$

TU/e

So why G-CNNs?

- Robustness.
- Better accuracy.
- Fewer parameters.
- Less data needed/no data augmentation needed.
- Reduced training time (measured in epochs at least).

The Homogeneous Space of Positions and Orientations

The Homogeneous Space of Positions and Orientations $\mathbb{M}^d$

- Defined as $\mathbb{M}^d = \mathbb{R}^d \times S^{d-1} \subseteq \mathbb{R}^d \times \mathbb{R}^d$
- The Lie group of translation and rotations $SE(d) = \mathbb{R}^d \rtimes SO(d)$ acts on it in the obvious way. We write the action with $\triangleright$.
- In the two-dimensional case there is only 1 group element that maps some reference point $p_0 \in \mathbb{M}^2$ onto any other point, i.e. $\mathbb{M}^2 = \mathrm{SE}(2)$.
- On $\mathbb{M}^2$ and $SE(2)$ we identify points with the triplet $(x, y, \theta) \in \mathbb{R}^2 \times \mathbb{R}/(2\pi\mathbb{Z})$.
- We will be focusing on the two-dimensional case.

Left-invariant Metric on the Space of Positions and Orientations

- We introduce the following left-invariant metric on $\mathbb{M}^d$.
- This turns the space into a Riemannian manifold, and we define the distance between two points the standard way.
- The ratio between w_m and w_l is called the anisotropy of the metric.
- If we let $w_l \to \infty$ we converge to a sub-Riemannian manifold.
- Why left-invariant…?

$$p = (x, n) \in \mathbb{M}^d \qquad \dot{p} = (\dot{x}, \dot{n}) \in T_p\mathbb{M}^d$$

$$\mathcal{G}_p(\dot{p}, \dot{p}) = w_m^2 |\dot{x} \cdot n|^2 + w_l^2 \|\dot{x} \wedge n\|^2 + w_a^2 \|\dot{n}\|^2$$

$$d(p, q) = \inf_{\substack{\gamma(0)=p, \\ \gamma(1)=q}} \int_0^1 \|\dot{\gamma}(t)\| dt$$

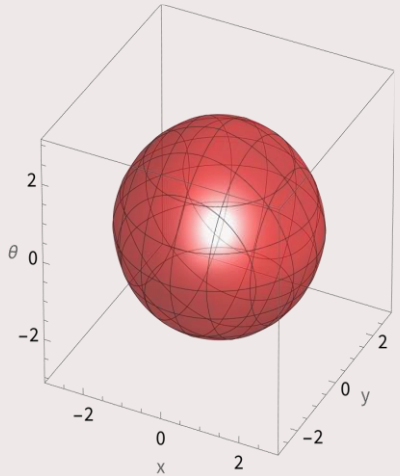
$$d(g \triangleright p, g \triangleright q) = d(p, q)$$

TU/e

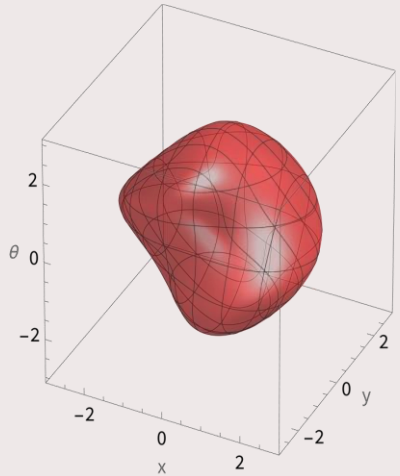
Visualization of the Metric on $\mathbb{M}^2$

- We can visualize the metric by plotting the Riemannian balls.

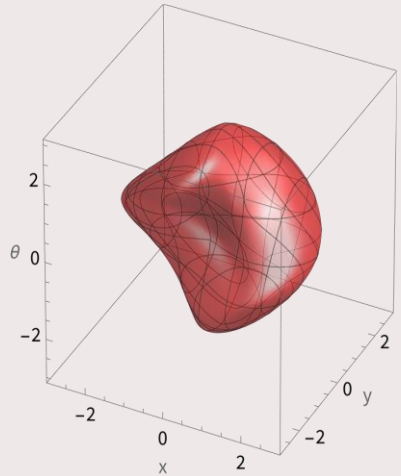
$$B_r = \{p \in \mathbb{M}^2 \mid d_{\mathcal{G}}(p_0, p) \leq r\}$$



$w_m = 1, w_l = 1, w_a = 1$



$w_m = 1, w_l = 2, w_a = 1$



$w_m = 1, w_l = 4, w_a = 1$

PDE-G-CNNs

Advection/Convection

- Advection is defined by the following PDE.
- Takes care of transport/offsetting.
- If the vector field c is left-invariant advection is equivariant.

$$\begin{cases} \dfrac{\partial W}{\partial t} = -cW \\ W\Big|_{t=0} = U \end{cases}$$

Diffusion

- Is defined by the following PDE.
- Regularizes.
- Solved by group convolution $*$ with the diffusion kernel K_t^α.
- The diffusion kernel can be well-approximated by a function of only distance from origin p_0
- If the metric is left-invariant diffusion is equivariant.

$$\begin{cases} \frac{\partial W}{\partial t} = \Delta_{\mathcal{G}} W \\ W\big|_{t=0} = U \end{cases}$$

$$W = K_t * U$$

$$(k * f)(p) \coloneqq \int_G k(g^{-1} \triangleright p) f(g \triangleright p_0) dg$$

$$K_t(p) \approx \mu_t \exp\left(-\frac{d(p_0, p)^2}{4t}\right)$$

Dilation/Erosion

- Is defined by the following PDE.
- Is max/min pooling and activation.
- Is solved by morphological group convolution $\square$ with morphological kernel k_t^α.
- Kernel is only a function of distance.
- If the metric is left-invariant dilation/erosion is equivariant.

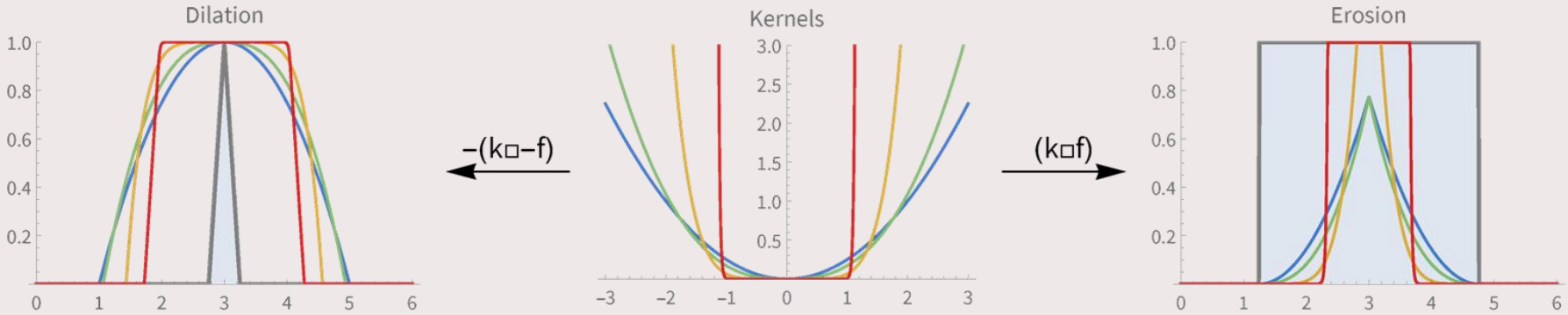
$$\begin{cases} \frac{\partial W}{\partial t} = -\frac{1}{\alpha}\left\|\nabla_{\mathcal{G}} W\right\|^\alpha \\ W\Big|_{t=0} = U \end{cases} \qquad \begin{cases} \frac{\partial W}{\partial t} = +\frac{1}{\alpha}\left\|\nabla_{\mathcal{G}} W\right\|^\alpha \\ W\Big|_{t=0} = U \end{cases}$$

$$W = k_t^\alpha \square U \qquad W = -(k_t^\alpha \square - U)$$

$$(k \square f)(p) \coloneqq \inf_{g \in G} k(g^{-1} \triangleright p) + f(g \triangleright p_0)$$

$$k_t^\alpha(p) \coloneqq \frac{t}{\beta}\left(\frac{d(p_0, p)}{t}\right)^\beta \qquad \frac{1}{\alpha} + \frac{1}{\beta} = 1$$

Visualization of Dilation/Erosion on $\mathbb{R}$



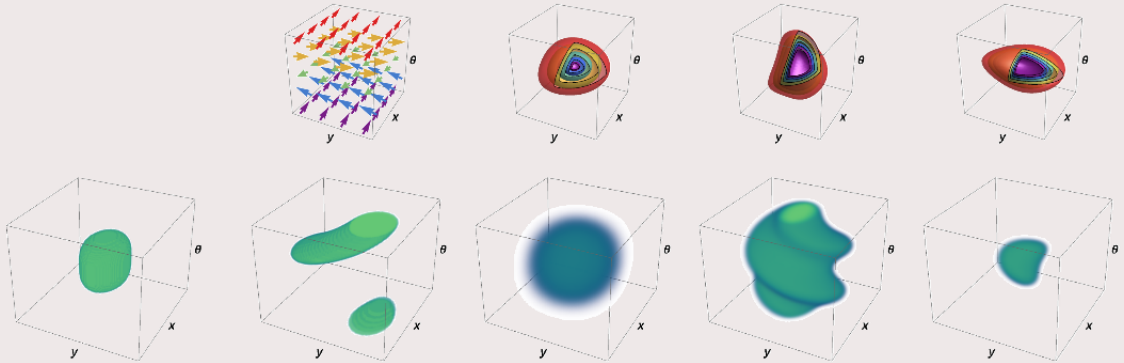
Visualization of the PDEs on $\mathbb{R}^2$

$\mathbb{R}^2$	Transport	Regularization	Max pooling	Min pooling
PDE Term:	$-\boldsymbol{c}W$	$-(-\Delta_{\mathcal{G}_1})^{\alpha}W$	$+\left\|\nabla_{\mathcal{G}_2^+}W\right\|_{\mathcal{G}_2^+}^{2\alpha}$	$-\left\|\nabla_{\mathcal{G}_2^-}W\right\|_{\mathcal{G}_2^-}^{2\alpha}$
Parameters:	transport vector $\begin{pmatrix}100\\120\end{pmatrix}$	Riemannian metric tensor $\begin{pmatrix}1&0\\0&1\end{pmatrix}$	Riemannian metric tensor $\begin{pmatrix}9&0\\0&1\end{pmatrix}$	Riemannian metric tensor $\begin{pmatrix}5&11\\11&40\end{pmatrix}$
Operation:	resample with offset	convolution with kernel	dilation with kernel	erosion with kernel

TU/e

Visualization of the PDEs on $\mathbb{M}^2$

$\mathbb{M}_2$	Transport	Regularization	Max pooling	Min pooling
PDE Term:	$-\boldsymbol{c}W$	$-(-\Delta_{\mathcal{G}_1})^{\alpha}W$	$+\left\|\nabla_{\mathcal{G}_2^+}W\right\|^{2\alpha}_{\mathcal{G}_2^+}$	$-\left\|\nabla_{\mathcal{G}_2^-}W\right\|^{2\alpha}_{\mathcal{G}_2^-}$
Parameters:	G–invariant vector field $\begin{pmatrix}1\\0\\2\end{pmatrix}$	Riemannian metric tensor $\begin{pmatrix}4&0&0\\0&6&0\\0&0&1\end{pmatrix}$	Riemannian metric tensor $\begin{pmatrix}4&0&0\\0&9&0\\0&0&0.5\end{pmatrix}$	Riemannian metric tensor $\begin{pmatrix}6&0&0\\0&3&0\\0&0&1.5\end{pmatrix}$
Operation:	resample with offset	convolution with kernel	dilation with kernel	erosion with kernel



PDE-G-CNNs

- The PDEs under consideration are all equivariant.
- This motivates the creation of so-called PDE-G-CNNs.
- Introduced by B.M.N. Smets et al. in the 2021 paper “PDE-based Group Equivariant Convolutional Neural Networks”.
- The parameters of PDE-G-CNNs are the metric parameters (w_m, w_l, w_a)

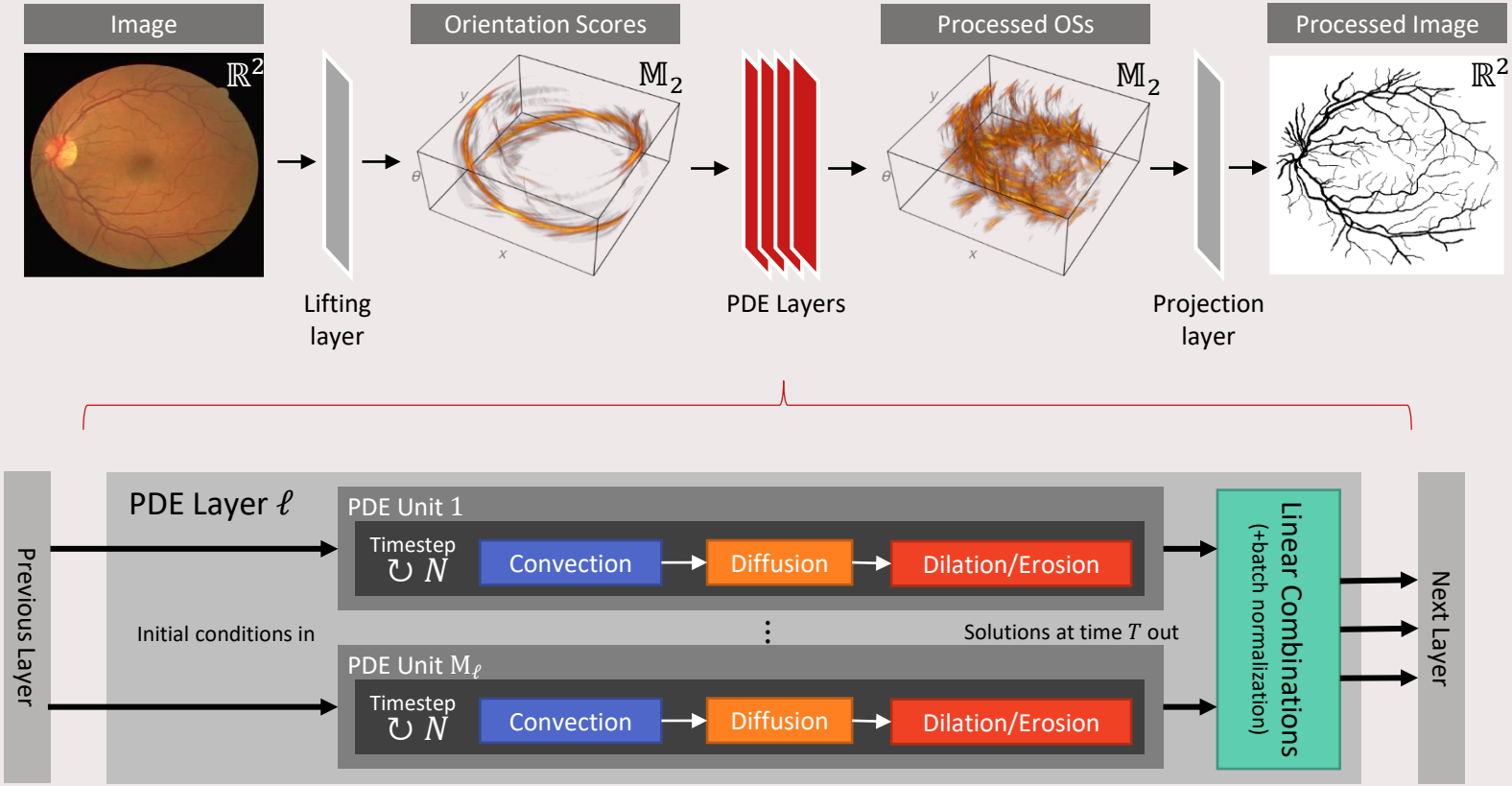
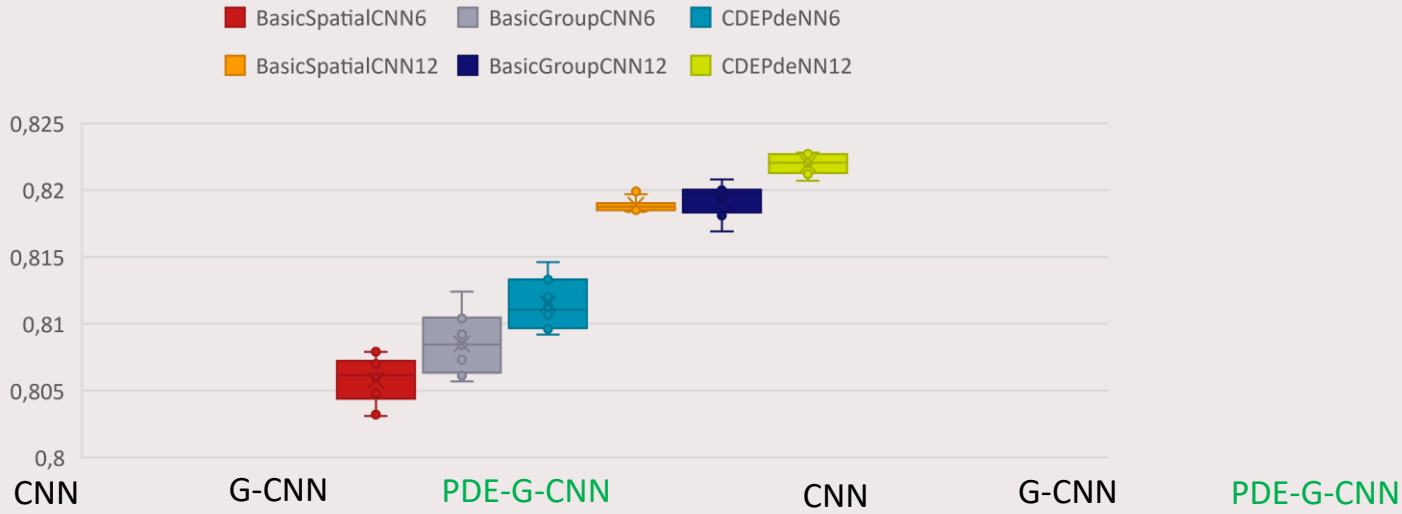


Image
$\mathbb{R}^2$
Lifting layer
Orientation Scores
$\mathbb{M}_2$
PDE Layers
Processed OSs
$\mathbb{M}_2$
Projection layer
Processed Image
$\mathbb{R}^2$
Previous Layer
PDE Layer ℓ
PDE Unit 1
Timestep
$\circlearrowleft N$
Convection
Diffusion
Dilation/Erosion
Initial conditions in
Solutions at time T out
PDE Unit M_ℓ
Timestep
$\circlearrowleft N$
Convection
Diffusion
Dilation/Erosion
Linear Combinations
(+batch normalization)
Next Layer

Results of PDE-G-CNNs in Vessel Segmentation



Parameters:	**47352**	**39258**	**4128**	**129432**	**114378**	**3678**
Layers:	6	6	6	12	12	12
LR	0.01	0.01	0.01	0.01	0.01	0.01
LR gamma	0.95	0.95	0.95	0.96	0.96	0.96
Epochs	60	60	60	80	80	80
Epoch time (s)	3	74	35	6	221	49
Test time (s):	**1.7**	**6.5**	**6.8**	**2.2**	**14.1**	**9.8**
6->12 slowdown				29%	117%	44%
Equivariance	T	RT	RT	T	RT	RT

So Why PDE-G-CNNs?

- All the benefits of G-CNNs
- Considerably less parameters.
- Geometric interpretable parameters.
- No activation functions needed.
- No loss in accuracy.

The Logarithmic Approximation

Logarithmic Distance Approximation

- M homogeneous space with Lie group G and left-invariant metric $\mathcal{G}$.
  Pick $p_0 \in M$, define projection $\pi(g) \coloneqq g \triangleright p_0$.
  $d(p) \leq \|\pi_* \log g\|_{\mathcal{G}}$ for any $g \in G$ such that $g \triangleright p_0 = p$.
- Therefor logarithmic distance approximation:
  $\rho(p) \coloneqq \inf_{g \in G_p} \|\pi_* \log g\|_{\mathcal{G}}$.
- One has $d \leq \rho$. But how good is the approximation really?
- We approximate kernels:

$$k_t^{\alpha}(p) \coloneqq \frac{t}{\beta}\left(\frac{d}{t}\right)^{\beta} \longrightarrow \bar{k}_t^{\alpha}(p) \coloneqq \frac{t}{\beta}\left(\frac{\rho}{t}\right)^{\beta}$$

Logarithmic Coordinates on $SE(2)$

- Let $g = (x, y, \theta) \in \mathbb{R}^2 \times [-\pi, \pi)$, then its logarithm $\log g$ is…
- Logarithmic distance approximation on $\mathbb{M}^2$ becomes…

$$\log g = \log(x, y, \theta) = c^1 A_1 + c^2 A_2 + c^3 A_3 \in T_e G$$

$$c^1 = \left(x \cos\left(\frac{\theta}{2}\right) + y \sin\left(\frac{\theta}{2}\right) \right) / \operatorname{sinc}\left(\frac{\theta}{2}\right)$$

$$c^2 = \left(-x \sin\left(\frac{\theta}{2}\right) + y \cos\left(\frac{\theta}{2}\right) \right) / \operatorname{sinc}\left(\frac{\theta}{2}\right)$$

$$c^3 = \theta$$

$$\rho := \sqrt{(w_m c^1)^2 + (w_l c^2)^2 + (w_a c^3)^2}$$

How good are the Approximations?

Symmetries of the (Approximative) Distance on $\mathbb{M}^2$

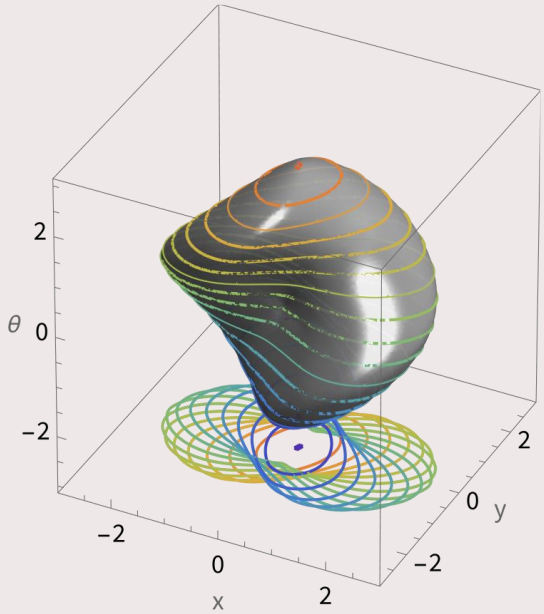
Theorem 1: *The exact and approximative (sub)-Riemannian distances are invariant under all ε^i symmetries*

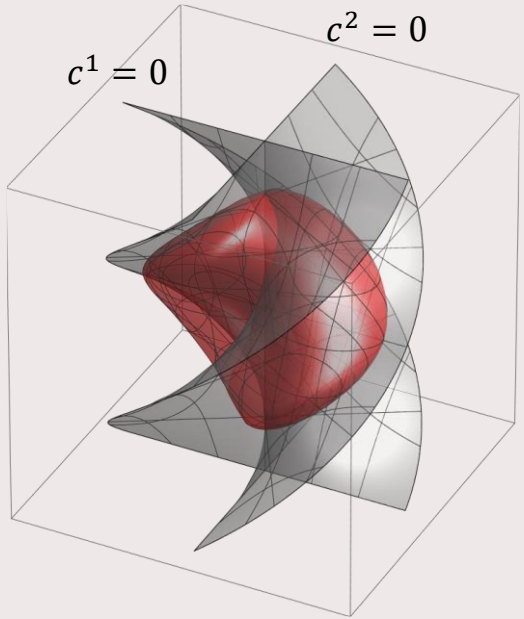
- In logarithmic coordinates these symmetries become more transparent.
- ρ has the correct global symmetries!

$$\rho := \sqrt{(w_m c^1)^2 + (w_l c^2)^2 + (w_a c^3)^2}$$

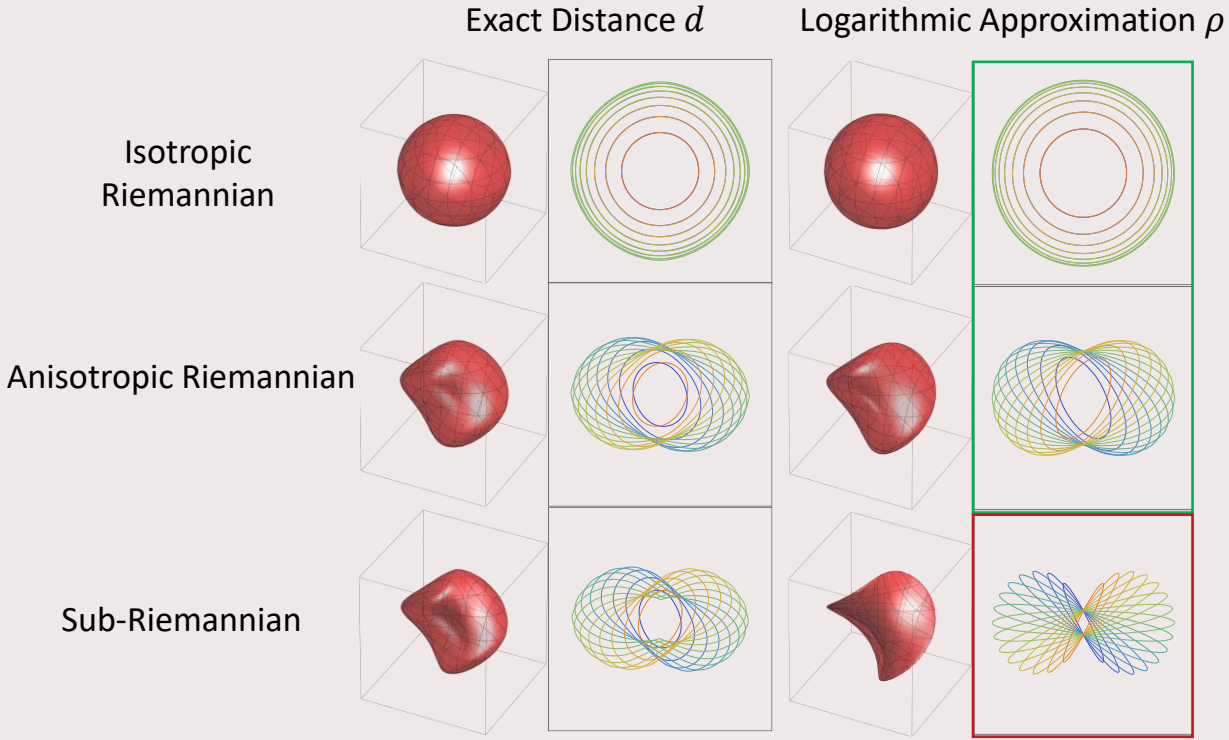
$$\begin{aligned}
\varepsilon^0(c_1, c_2, c_3) &= (c_1, c_2, c_3) \\
\varepsilon^1(c_1, c_2, c_3) &= (c_1, -c_2, c_3) \\
\varepsilon^2(c_1, c_2, c_3) &= (-c_1, c_2, c_3) \\
\varepsilon^3(c_1, c_2, c_3) &= (-c_1, -c_2, c_3) \\
\varepsilon^4(c_1, c_2, c_3) &= (-c_1, c_2, -c_3) \\
\varepsilon^5(c_1, c_2, c_3) &= (-c_1, -c_2, -c_3) \\
\varepsilon^6(c_1, c_2, c_3) &= (c_1, c_2, -c_3) \\
\varepsilon^7(c_1, c_2, c_3) &= (c_1, -c_2, -c_3)
\end{aligned}$$

Visualization of the Symmetries





Exact and Approximative Balls



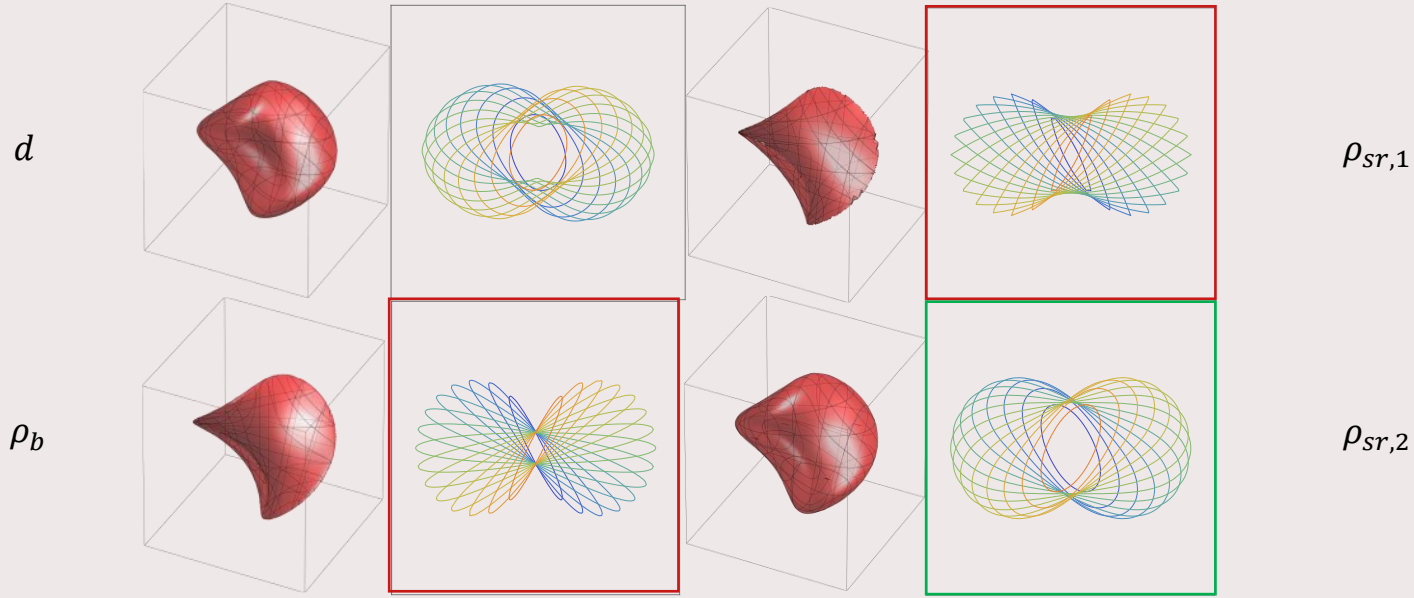
Sub-Riemannian Approximations

- To solve the bad behavior at high anisotropy we introduce two sub-Riemannian approximations

$$\rho_{sr,1} \coloneqq \sqrt{8 w_m w_a |c^2| + (w_m c^1)^2 + (w_a c^3)}$$

$$\rho_{sr,2} \coloneqq \sqrt[4]{\beta (w_m w_a c^2)^2 + \left((w_m c^1)^2 + (w_a c^3)\right)^2}$$

Exact and Approximative Balls Sub-Riemannian



Quantitative Local Analysis

Theorem 2: locally around p_0 we have that

$$\rho^2(1-\varepsilon) \leq d^2 \leq \rho^2$$

$$\bar{k}_t^\alpha \geq k_t^\alpha \geq \bar{k}_t^\alpha \left(1 - \frac{\alpha\varepsilon}{2\alpha - 1} + \mathcal{O}(\varepsilon^2)\right)$$

where

$$\varepsilon \leq \rho^2 \left(\frac{\left|w_m^2 - w_l^2\right| M + M^2}{12 w_m^2 w_l^2 w_a^2} + \mathcal{O}(\rho^2)\right)$$

$$M = \max\{w_m^2, w_l^2\}$$

- For high anisotropy the error explodes.
- For $w_a \to 0$ the error explodes

- For $w_a \to \infty$ the error goes to 0
- For low anisotropy and $w_a \gg w_l$ **large accurate regions**

Conclusion

Conclusion

- PDE-G-CNNs better than G-CNNs, which are better than CNNs.

- We have assessed the quality of the approximations used:
  - The approximations have all the symmetries the exact distances have.
  - Assessing the quality of the PDE-G-CNNs boils down to the quality of the approximative (sub)Riemannian distance.
  - We found by asymptotic analysis of the relative error between the true and approximative distance conditions on the parameters.
  - This improves upon results by Smets et al.

- First milestone almost done: JMIV paper and conference paper SSVM.

Further Research

- Extension of PDE-G-CNN's
  - It is claimed our vision carries a Sub-Riemannian structure.
  - Is the geometry of our vision the geometry for artificial vision?
  - Riemannian vs. Sub-Riemannian PDE-G-CNNs?
  - Inclusion of curvature (non-diagonal metric tensor field).
- Axiomatic approach for geometric deep learning.
  - There are some preliminary ideas that force our choice of PDEs.

Questions?