On the topological and geometrical properties of group-equivariant non-expansive operators and their use in Deep Learning

Nicola Quercioli

February 9, 2021

The content of this talk

In this talk we will illustrate the topological theory introduced in this paper:

Mattia G. Bergomi, Patrizio Frosini, Daniela Giorgi, Nicola Quercioli, *Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning*, Nature Machine Intelligence, vol. 1, n. 9, pages 423–433 (2 September 2019).

Full-text access to a view-only version of this paper is available at the link `https://rdcu.be/bP6HV`.

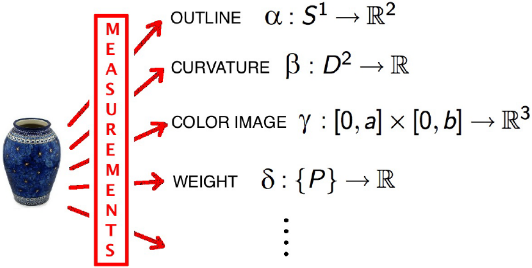
Outline

The role of equivariant operators in machine learning

- As pointed out by several authors (Mallat, Poggio, Rosasco...) the role of equivariant operators in machine learning is getting more and more important.
- Data comparison is almost never a direct process: it is usually mediated by the comparison of new data produced by agents/observers, i.e. by the comparison of the action of agents on the original data.
- Equivariant operators (and networks of equivariant operators) can model agents acting on data.

What does MEASUREMENT mean?

Measurement is the assignment of a number to a characteristic of an object or event, which can be compared with other objects or events.



Assumptions in our model

- Data are represented as functions defined on topological spaces, since only data that are stable w.r.t. a certain criterion (e.g., with respect to some kind of measurement) can be considered for applications, and stability requires a topological structure.
- Data cannot be studied in a direct and absolute way. They are only knowable through acts of transformation made by an agent/observer. From the point of view of data analysis, only the pair (data, agent) matters. In general terms, agents are not endowed with purposes or goals: they are just ways and methods to transform data. Acts of measurement are a particular class of acts of transformation.
- Agents are described by the way they transform data while respecting some kind of invariance. In other words, any agent can be seen as a group equivariant operator acting on a function space.
- Data similarity depends on the output of the considered agent.

The space of admissible measurements

First, for us a data set is given by a set of bounded real-valued measurements on some set X:

$$\Phi = \{\varphi \colon X \to \mathbb{R}\} \subseteq \mathbb{R}_b^X,$$

where $\mathbb{R}_b^X$ is the set of all bounded real-valued functions on X.

The role of topology in our framework

If we wish to develop a theory that can be applied in real situations, we need stability with respect to noise.

A natural topology on the set Φ of possible measurements is the one induced by the L^∞-metric

$$D_\Phi(\varphi_1, \varphi_2) := \|\varphi_1 - \varphi_2\|_\infty = \sup_{x \in X} |\varphi_1(x) - \varphi_2(x)|. \tag{1}$$

Example

- $X := S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$,
- $\Phi =$ set of all non-expansive functions from S^1 to $[0, 1]$.

A topology on the space X

Since measurements are the central concept in our approach, the topology on X is derived from D_Φ.

We define an extended pseudo-metric D_X on X by setting

$$D_X(x_1, x_2) := \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|. \tag{2}$$

Proposition

Every function in Φ is continuous.
If Φ is totally bounded, the topology induced by D_X is the coarsest topology such that the elements of Φ are continuous.

Φ-preserving isometries with respect to D_X

Set of Φ-preserving bijections:

$$\mathrm{Aut}_\Phi(X) := \{g\colon X \to X,\ \varphi g, \varphi g^{-1} \in \Phi \text{ for every } \varphi \in \Phi\}. \tag{3}$$

Theorem

$$\mathrm{Aut}_\Phi(X) = \mathrm{Iso}_\Phi(X)$$

We can define an extended pseudo-metric

$$D_{\mathrm{Aut}}(g_1, g_2) := \sup_{x \in X} D_X(g_1(x), g_2(x)) = \sup_{\varphi \in \Phi} D_\Phi(\varphi g_1, \varphi g_2). \tag{4}$$

- $\mathrm{Aut}_\Phi(X)$ *is a topological group;*
- *the action of* $\mathrm{Aut}_\Phi(X)$ *on* Φ *through right composition is continuous.*

We usually consider a group $G \subseteq \mathrm{Aut}_\Phi(X)$ as an invariance group.
We will say that (Φ, G) is a PERCEPTION PAIR.

Examples:

Consider a perception pair (Φ, G), where:

- $\Phi =$ non-expansive compactly supported functions from $X = \mathbb{R}^2$ to a suitable interval $[a, b]$;
- G is the group of isometries of $\mathbb{R}^2$

Consider a perception pair (Ψ, H), where:

- $Y := S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$,
- $\Psi =$ set of all non-expansive functions from S^1 to $[0, 1]$,
- $H =$ group of rotations of S^1.

Our ground truth: the natural pseudo-distance d_G

Definition

Suppose that G is a group of Φ-preserving self-homeomorphisms of X. The pseudo-distance $d_G : \Phi \times \Phi \to \mathbb{R}$ is defined by setting

$$d_G(\varphi_1, \varphi_2) = \inf_{g \in G} D_\Phi(\varphi_1, \varphi_2 g). \tag{5}$$

It is called the natural pseudo-distance associated with the group G.

If G_1 and G_2 are groups and $G_1 \subseteq G_2 \subseteq \mathrm{Homeo}_\Phi(X)$, then the definition of d_G implies that

$$d_{G_2}(\varphi_1, \varphi_2) \leq d_{G_1}(\varphi_1, \varphi_2) \leq D_\Phi(\varphi_1, \varphi_2) \tag{6}$$

for every $\varphi_1,\ \varphi_2 \in \Phi$.

The natural pseudo-distance d_G

The natural pseudo-distance d_G represents our ground truth.

Unfortunately, in many cases d_G is difficult to compute.

Remark

We can easily find groups $G \subseteq \mathrm{Homeo}_\Phi(X)$ that cannot be approximated with arbitrary precision by smaller finite subgroups.

Nevertheless, in this talk we will show that d_G can be approximated with arbitrary precision by means of a **DUAL** approach based on persistent homology and group equivariant non-expansive operators (GENEOs).

The space of GENEOs

Definition

Assume that (Φ, G), (Ψ, H) are two perception pairs. A *Group Equivariant Non-Expansive Operator (GENEO)* is a map $(F, T) : (\Phi, G) \to (\Psi, H)$ such that the following properties hold for every $\varphi_1, \varphi_2 \in \Phi$:

- $T : G \to H$ is a group homomorphism;
- $F(\varphi_1 g) = F(\varphi_1) T(g)$ for every $g \in G$;
- $D_\Psi (F(\varphi_1), F(\varphi_2)) \leq D_\Phi (\varphi_1, \varphi_2)$.

If F_1 and F_2 are two GENEOs, we set

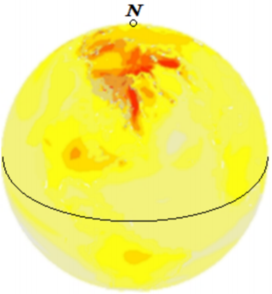
$$D_{\mathrm{GENEO}} (F_1, F_2) := \sup_{\varphi \in \Phi} D_\Psi (F_1(\varphi), F_2(\varphi)) \tag{7}$$

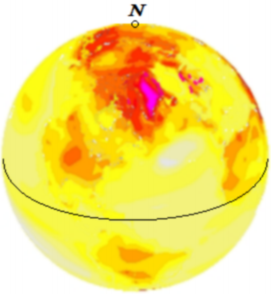
D_{GENEO} is a metric on the space of all GENEOs.

We will use the symbol $\mathcal{F}_T^{\mathrm{all}}$ to denote the set of all GENEOs from (Φ, G) to (Ψ, H) with respect to T.

An example of GENEO

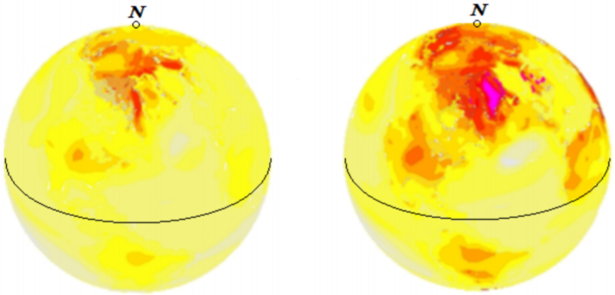
We give an example of the use of the definition of GENEO between two different perception pairs (Φ, G), (Ψ, H).





An example of GENEO

- $X = S^2$
- $\Phi =$ set of non-expansive functions from S^2 to a fixed interval $[a, b]$
- $G =$ group of rotations of S^2 around the axis $N - S$



- $Y =$ the equator S^1 of S^2
- $\Psi =$ set of non-expansive functions from S^1 to $[a, b]$
- $H =$ group of rotations of S^1

An example of GENEO

A simple example of GENEO from (Φ, G) to (Ψ, H) is built by setting

- $T(g)$ equal to the rotation $h \in H$ of the equator S^1 that is induced by the rotation g of S^2 for every $g \in G$.
- $F(\varphi)(\theta) = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \varphi(\theta, \alpha)\, d\alpha$ for every $\varphi \in \Phi$.

Another example of GENEO

Perception pair

- $\Phi =$ non-expansive compactly supported functions from $X = \mathbb{R}^2$ to a suitable interval $[a, b]$;
- G is the group of isometries of $\mathbb{R}^2$

A simple example of GENEO from (Φ, G) to itself with respect to Id:

- Take a compactly supported integrable function $f : \mathbb{R}^2 \to \mathbb{R}$ that is invariant with respect to every rotation around the point $(0, 0)$ and suppose that $\int_{\mathbb{R}^2} f(x, y)dxdy = 1$.
- $F(\varphi) := f * \varphi$ is a group equivariant non-expansive operator with respect to the homomorphism $\mathrm{id}_G : G \to G$.

Some good properties

Theorem

If Φ and Ψ are compact, $\mathcal{F}_T^{\mathrm{all}}$ is compact *with respect to* D_{GENEO}.

Corollary

$\mathcal{F}_T^{\mathrm{all}}$ *can be ϵ-approximated by a finite subset.*

Theorem

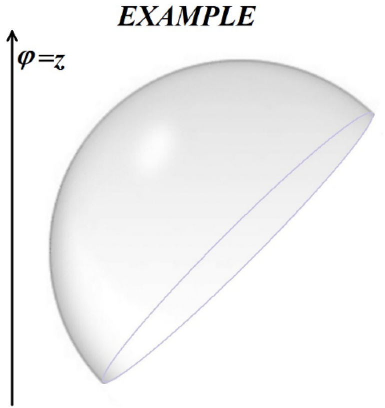
If Ψ is convex, then $\mathcal{F}_T^{\mathrm{all}}$ **is convex**.

What is persistent homology?

What is persistent homology?

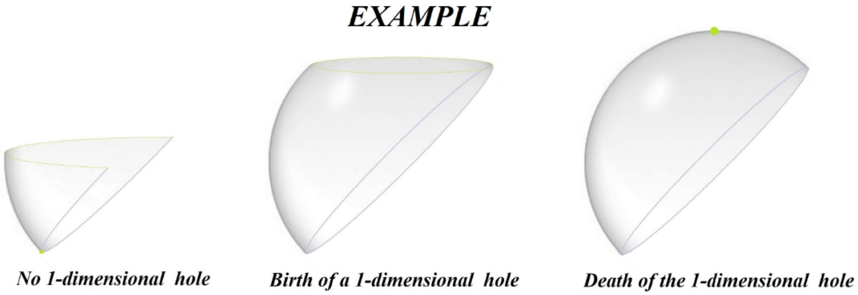
What is persistent homology?

If $\varphi : X \longrightarrow \mathbb{R}$ is a continuous functions, we can consider the sublevel sets $X_t := \{x \in X : \varphi(x) \leq t\}$. When t varies we see the birth and death of k-dimensional holes.



What is persistent homology?

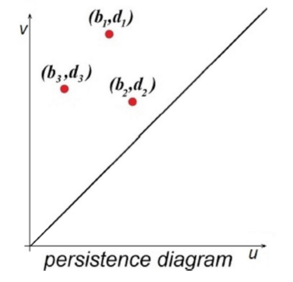
If $\varphi : X \longrightarrow \mathbb{R}$ is a continuous functions, we can consider the sublevel sets $X_t := \{x \in X : \varphi(x) \leq t\}$. When t varies we see the birth and death of k-dimensional holes.



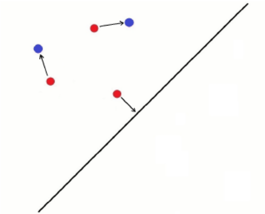
What is persistent homology?

Persistence diagram in degree k of φ

The multiset of the pairs (b_i, d_i) where b_i and d_i are the times of birth and death of the i-th hole of dimension k.



Comparison of persistent diagrams



Persistence diagrams can be compared by means of the bottleneck distance. Let $\mathrm{Dgm}_k(\varphi_1), \mathrm{Dgm}_k(\varphi_2)$ be two persistence diagrams.

$$d_{\text{match}}(\mathrm{Dgm}_k(\varphi_1), \mathrm{Dgm}_k(\varphi_2)) := \inf_{\sigma} \sup_{x \in \mathrm{Dgm}_k(\varphi_1)} d^*(x, \sigma(x)), \tag{8}$$

where $\sigma : \mathrm{Dgm}_k(\varphi_1) \to \mathrm{Dgm}_k(\varphi_2)$ is a bijection.

A link with TDA - 1

Persistent homology enters this theoretical framework by means of an equality allowing us to approximate the natural pseudo-distance:

Theorem

Assume $(\Phi, G) = (\Psi, H)$, every function in Φ is non-negative, the k-th Betti number of X does not vanish, and Φ contains each constant function c for which a function $\varphi \in \Phi$ exists such that $0 \leq c \leq \|\varphi\|_\infty$. Then

$$d_G(\varphi_1, \varphi_2) = \sup_{F \in \mathcal{F}^{\mathrm{all}}} d_{\mathrm{match}}\left(\mathrm{Dgm}_k(F(\varphi_1)), \mathrm{Dgm}_k(F(\varphi_2))\right) \tag{9}$$

where $\mathrm{Dgm}_k(F(\varphi_i))$ is the persistence diagram in degree k of the function $F(\varphi_i)$ and d_{match} is the usual bottleneck distance.

As a consequence, the construction of new classes of GENEOs is an important step in the computation of the natural pseudo-distance. This fact justifies the interest in the following results.

A link with TDA - 1

Let us take a finite ε-approximation $\mathcal{F}$ of $\mathcal{F}_T^{\mathrm{all}}$. We can then define the pseudo-metric

$$D_{\mathrm{match}}^{\mathcal{F}}(\varphi_1, \varphi_2) := \sup_{F \in \mathcal{F}} d_{\mathrm{match}}\left(\mathrm{Dgm}_k(F(\varphi_1)), \mathrm{Dgm}_k(F(\varphi_2))\right) \tag{10}$$

where d_{match} is the usual bottleneck distance between persistence diagrams.

Properties

For any $\varphi_1,\ \varphi_2 \in \Phi$ and any $g \in G$:

- $D_{\mathrm{match}}^{\mathcal{F}}(\varphi_1, \varphi_2 g) = D_{\mathrm{match}}^{\mathcal{F}}(\varphi_1, \varphi_2)$;
- $D_{\mathrm{match}}^{\mathcal{F}}(\varphi_1, \varphi_2) \leq \|\varphi_1 - \varphi_2\|_\infty$;
- $|D_{\mathrm{match}}^{\mathcal{F}}(\varphi_1, \varphi_2) - d_G(\varphi_1, \varphi_2)| \leq 2\varepsilon$.

A link with TDA - 2

A proxy for the fast comparison of GENEOs:

$$\Delta_{\mathsf{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} d_{\mathrm{match}} \left(\mathrm{Dgm}_k(F_1(\varphi)), \mathrm{Dgm}_k(F_2(\varphi))\right). \tag{11}$$

Proposition

$$\Delta_{GENEO}(F_1, F_2) \leq D_{GENEO}(F_1, F_2). \tag{12}$$

In other words, persistent homology provides an efficient way for the comparison of GENEOs.

Building new GENEOs

How could we build new GENEOs from other GENEOs?

A first simple example:

Composition

Assume that

- F_1 is a GENEO from (Φ_1, G_1) to (Φ_2, G_2) with respect to $T_1 : G_1 \to G_2$
- F_2 is a GENEO from (Φ_2, G_2) to (Φ_3, G_3) with respect to $T_2 : G_2 \to G_3$

Then $F_2 \circ F_1$ is a GENEO from (Φ_1, G_1) to (Φ_3, G_3) with respect to $T_2 \circ T_1 : G_1 \to G_3$.

Building new GENEOs via translations, the maximum operator and weighted averages

Assume that two perception pairs (Φ, G), (Ψ, H) and a homomorphism $T : G \to H$ are given.
Assume also that $F_1, \ldots, F_n$ are GENEOs from (Φ, G) to (Ψ, H) with respect to T.

Translations

The operator $F_b(\varphi) := \varphi - b$ is a GENEO from (Φ, G) to (Ψ, H) with respect to T, for every $b \in \mathbb{R}$ such that $F_b(\Phi) \subseteq \Psi$.

Maximum operator

The operator $F(\varphi) := \max_i F_i(\varphi)$ is a GENEO from (Φ, G) to (Ψ, H) with respect to T, provided that $F(\Phi) \subseteq \Psi$.

Building new GENEOs via weighted averages and power means

Weighted averages

If $(a_1, \ldots, a_n) \in \mathbb{R}^n$ and $\sum_{i=1}^n |a_i| \leq 1$, then the operator $F(\varphi) := \sum_{i=1}^n a_i F_i(\varphi)$ is a GENEO from (Φ, G) to (Ψ, H) with respect to T, provided that $F(\Phi) \subseteq \Psi$.

Power Means

Let us define the operator $M_p(F_1, \ldots, F_n) : \Phi \longrightarrow C_b^0(Y, \mathbb{R})$ by setting

$$M_p(F_1, \ldots, F_n)(\varphi)(x) := \left(\frac{1}{n} \sum_{i=1}^n |F_i(\varphi)(x)|^p \right)^{\frac{1}{p}}. \quad (13)$$

If $p \geq 1$ and $M_p(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$, $M_p(F_1, \ldots, F_n)$ is a GENEO from (Φ, G) to (Ψ, H) with respect to T.

Building GENEOs via non-expansive functions

Assume that two perception pairs (Φ, G), (Ψ, H) and a homomorphism $T : G \to H$ are given.
Assume also that $F_1, \ldots, F_n$ are GENEOs from (Φ, G) to (Ψ, H) with respect to T.

Let $\mathcal{L}$ be a non-expansive map from $\mathbb{R}^n$ to $\mathbb{R}$, where $\mathbb{R}^n$ is endowed with the norm $\|(x_1, \ldots, x_n)\|_\infty := \max_{1 \le i \le n} |x_i|$.
Let us define $\mathcal{L}^*(F_1, ..., F_n)$ by setting

$$\mathcal{L}^*(F_1, \ldots, F_n)(\varphi)(x) := \mathcal{L}(F_1(\varphi)(x), \ldots, F_n(\varphi)(x)). \tag{14}$$

If $\mathcal{L}^*(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$, then $\mathcal{L}^*(F_1, \ldots, F_n)$ is a GENEO from (Φ, G) to (Ψ, H) with respect to T.

Building GENEOs via Permutant Measures

A finite Borel signed measure μ on $\mathrm{Aut}_\Phi(X)$ is called a *permutant measure* with respect to G if μ is invariant under the conjugation action of G (i.e., $\mu(H) = \mu(gHg^{-1})$ for every $g \in G$).

Assume Φ is contained in a linear separable space. Let H be non-null measurable subset of $\mathrm{Aut}_\Phi(X)$ such that for every $g \in G$ it holds that $gHg^{-1} \subseteq H$. We consider the operator $F_H \colon \Phi \to \mathbb{R}_b^X$ defined as

$$F_H(\varphi)(x) := \frac{1}{\mu(H)} \int_H \varphi h \, d\mu(h).$$

Proposition

(F_H, id_G) is a GENEO from (Φ, G) to $(\mathbb{R}_b^X, G)$.

Representation of linear GENEOs

Let X be a finite set. Assume that $G \subseteq \mathrm{Aut}_{\mathbb{R}^X}(X)$ transitively acts on X.

Theorem

Consider a map $F\colon \mathbb{R}^X \to \mathbb{R}^X$.
The map F is a linear GENEO from $(\mathbb{R}^X, G)$ to itself with respect to id_G if and only if a permutant measure μ exists such that $F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} \varphi h^{-1} \, \mu(h)$ for every $\varphi \in \mathbb{R}^X$, and $\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| \leq 1$.

Metric Learning

Data set:

$\Phi = \{\varphi_i : X \to \mathbb{R}\}_{i=1}^{n}$ equipped with a labeling function $L : \Phi \to I = \{1, \ldots, n\}$.

Assumption:

Data labelled with the same symbol share common features w.r.t. the agent we want to approximate.

Goal:

To illustrate an algorithm for metric learning based on the metrics introduced on the space of GENEOs.

Operator selection and sampling

Let $\mathcal{F}$ be the space of operators that will act on the samples.

Sampling:

$\mathcal{C} = \{F_k\}_{k\in\{1,\dots,N\}}$ set of randomly sampled N candidate operators in $\mathcal{F}$.

The data set can be written as:

$$\Phi = \bigsqcup_{i\in I} \Phi_i \tag{15}$$

where $\Phi_i = L^{-1}(i)$.

Selection:

For each candidate $F \in \mathcal{C}$, we define the label-dependent value

$$s_\ell(F) := \max_{\varphi_i^\ell, \varphi_j^\ell \in \Phi_\ell} d_{\text{match}}(\mathrm{Dgm}_k(F(\varphi_i^l)), \mathrm{Dgm}_k(F(\varphi_j^l))) \tag{16}$$

F is selected if $s_\ell(F)$ is smaller than a fixed threshold ε for every ℓ.

Sampling / Elimination of redundant information

Given a class ℓ, we define the distance between two operators F_p and F_q

$$\Delta^{\ell}_{\mathrm{GENEO}}(F_p, F_q) := \sup_{\varphi_\ell \in \Phi_\ell} d_{\mathrm{match}}(\mathrm{Dgm}_k(F_p(\varphi_\ell)), \mathrm{Dgm}_k(F_q(\varphi_\ell)). \tag{17}$$

- For every label ℓ, we sort the pairs (F_p, F_q) in ascending order of $\Delta^{\ell}_{\mathrm{GENEO}}$, and assign to each pair of operators its index in the sorted list of distances.
- We then define the *interclass contrastive score* of the pair (F_p, F_q) as the sum of its indices over all classes.
- Finally, we remove redundant operators, i.e. we select only one operators for pairs whose score is below a fixed threshold t.

Finally, two objects φ_1 and φ_2 can be compared by computing the pseudo-metric $D^{\mathcal{S}}_{\mathrm{match}}(\varphi_1, \varphi_2)$, where $\mathcal{S}$ is the set of the selected and sampled operators.

An example of Metric Learning: the isometry case

Data set:

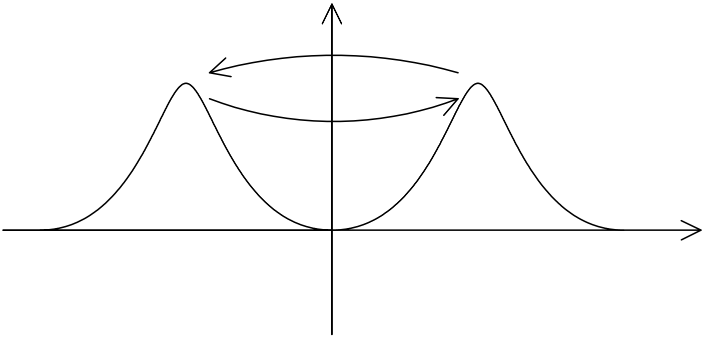
Φ = set of compactly supported continuous functions $\phi : \mathbb{R}^2 \to \mathbb{R}$

$$g_\tau(t) := e^{-\frac{(t-\tau)^2}{2\sigma^2}}$$

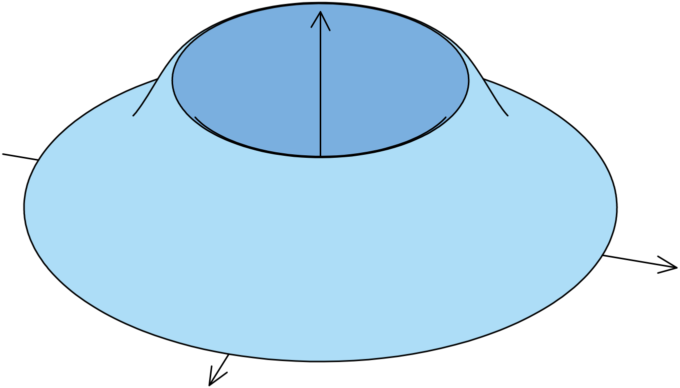
For each $p = (a_1, \tau_1, \ldots, a_k, \tau_k) \in R^{2k}$ with $\sum_{i=1}^k a_i^2 = \sum_{i=1}^k \tau_i^2 = 1$, we then consider the function $G_p : \mathbb{R}^2 \to \mathbb{R}$ defined as
$G_p(x, y) := \sum_{i=1}^k a_i g_{\tau_i}\left(\sqrt{x^2 + y^2}\right)$.

$$F_p(\phi)(x, y) := \int_{\mathbb{R}^2} \phi(\alpha, \beta) \cdot \frac{G_p(x - \alpha, y - \beta)}{\|G_p\|_{L^1}} \, d\alpha \, d\beta.$$

An example of Metric Learning: the isometry case



An example of Metric Learning: the isometry case



An example of Metric Learning: the isometry case

Data set:

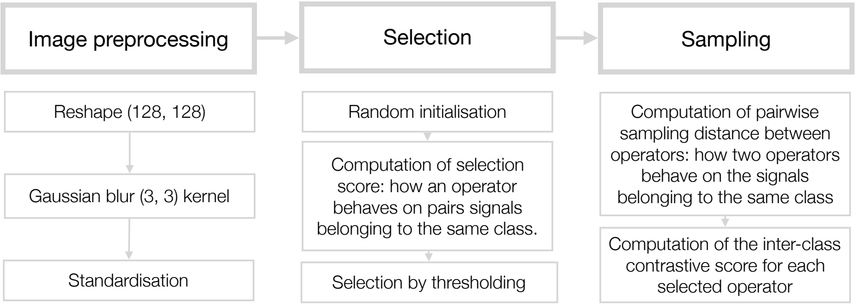
Φ = set of compactly supported continuous functions $\phi : \mathbb{R}^2 \to \mathbb{R}$

$$F_p(\phi)(x,y) := \int_{\mathbb{R}^2} \phi(\alpha,\beta) \cdot \frac{G_p(x-\alpha, y-\beta)}{\|G_p\|_{L^1}} \, d\alpha \, d\beta.$$

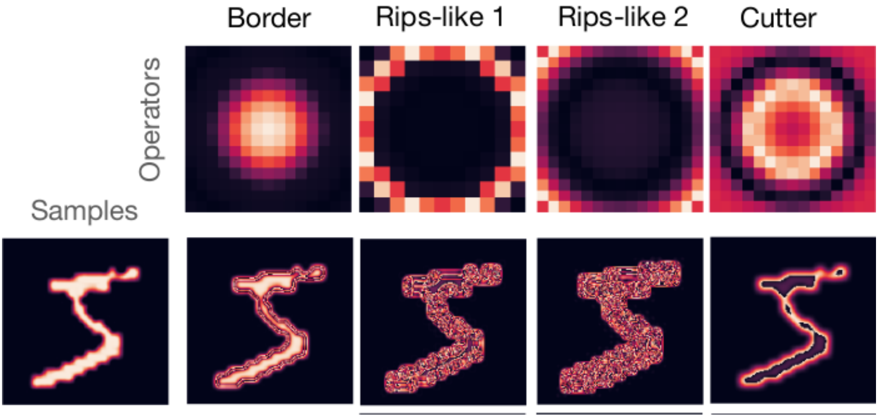
Operators that act on our data set:

We define a parametric family of IENEOs $\mathcal{F} = \{F_p\}_{p \in S}$ from Φ to $C^0_b(\mathbb{R}^2, \mathbb{R})$.

Experimental pipeline

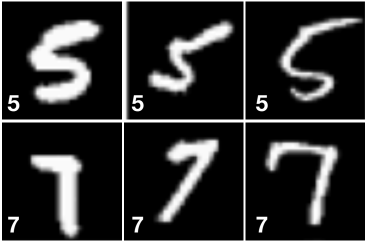


Isometry equivariant non-expansive operators on MNIST

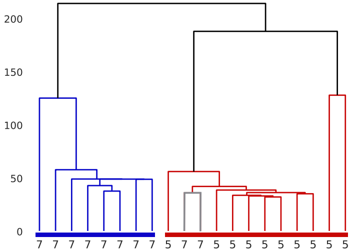


Metric Learning via GENEO selection and sampling

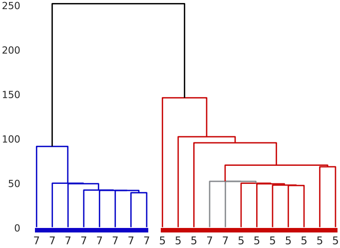
A. MNIST samples



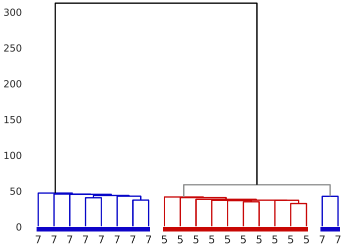
B. Metric on validation examples varying filter size. Here 7×7 px



C. Metric learned with filters of size 11×11 px

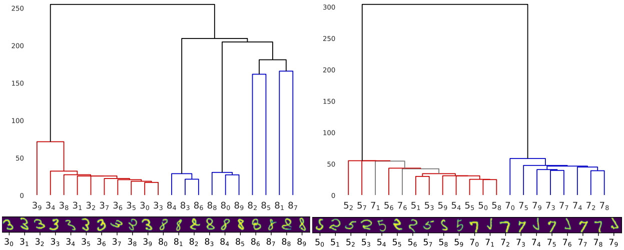


D. Metric learned with filters of size 21×21px

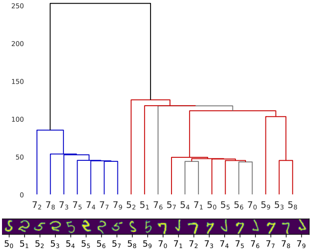


Metric Learning on augmented MNIST validation samples

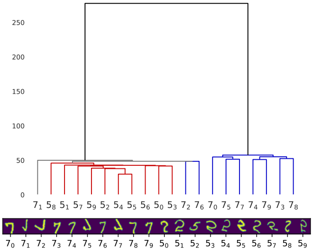
A. Metric learned with filters of size 7×7 px on selected and sampled IENEOs on classes (3,8) and (5,7), respectively



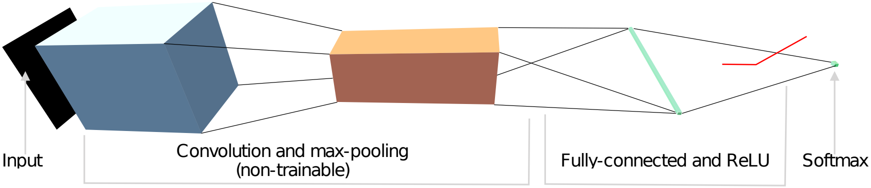
B. Metric learned with filters of size 11×11 px



B. Metric learned with filters of size 21×21 px



A first application: knowledge injection



Convolutional neural network architecture used in the knowledge injection experiment. The two first layers are a convolutional and maxpooling layer, respectively. The parameters of these layers are fixed. Two fully-connected layers counting 64 and 2 units respectively are trained to classify the two selected classes. The first fully-connected layer uses a rectified linear unit (ReLU) and the second a softmax as activation functions.

A first application: knowledge injection

Dataset	500 operators 20 examples	500 operators 40 examples	750 operators 20 examples	Random filters
MNIST	0.988591	0.963282	**0.991158**	0.506815
MNIST	**0.989732**	0.949096	0.987450	0.960501
MNIST	**0.990588**	0.960501	0.989447	0.506815
f-MNIST	0.984598	0.960223	**0.986309**	0.506815
f-MNIST	0.990017	0.958275	**0.990588**	0.958554
f-MNIST	**0.990302**	0.963561	0.988591	0.506815
CIFAR-10	0.982886	0.960223	**0.987735**	0.506815
CIFAR-10	**0.992299**	0.962170	0.988876	0.506815
CIFAR-10	0.982316	0.960501	**0.989732**	0.971627
MNIST $\mu(\sigma)$	0.989(0.001)	0.958(0.007)	0.989(0.001)	0.658(0.262)
f-MNIST $\mu(\sigma)$	0.988(0.003)	0.961(0.002)	0.988(0.002)	0.657(0.261)
CIFAR-10 $\mu(\sigma)$	0.985(0.005)	0.961(0.001)	0.988(0.001)	0.661(0.268)
All $\mu(\sigma)$	0.988(0.003)	0.959(0.004)	0.988(0.001)	0.659(0.228)

TAKE AWAY MESSAGE

1. Data are not so important.
2. The interactions between data and agents matter.
3. We should study the topological and geometric properties of the space of these interactions.

Conclusions and open questions

We have introduced a topological-geometrical model where we can formalize and attack this problem: how can we find efficient methods to approximate a given agent/observer by a GENEO belonging to a compact and convex space of GENEOs?

This problem leads us to the following open questions:

- How can we build a good library of GENEOs?
- How can we find a method to choose a finite set $\mathcal{F}^*$ of GENEOs that allows for both a good approximation of the natural pseudo-distance d_G and a fast computation?
- In which cases can the set of GENEOs be equipped with the structure of a Riemannian manifold?
- Could we compose operators to form networks, in the same way as computational units are connected in an artificial neural network?

Thanks for your attention!