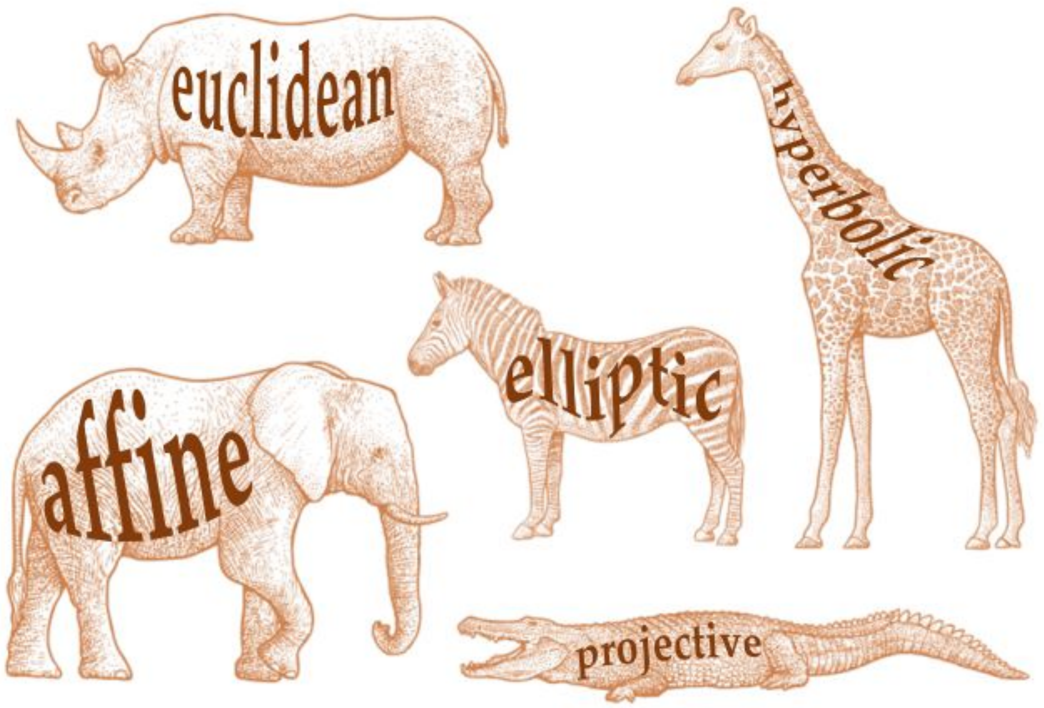
CENTRO DE CIENCIAS MATEMÁTICAS

Introducción a Geometric Deep Learning

Curso del PCCM - ML-GDL

¿Simetría?

“Zoológico” de geometrías en el siglo 19

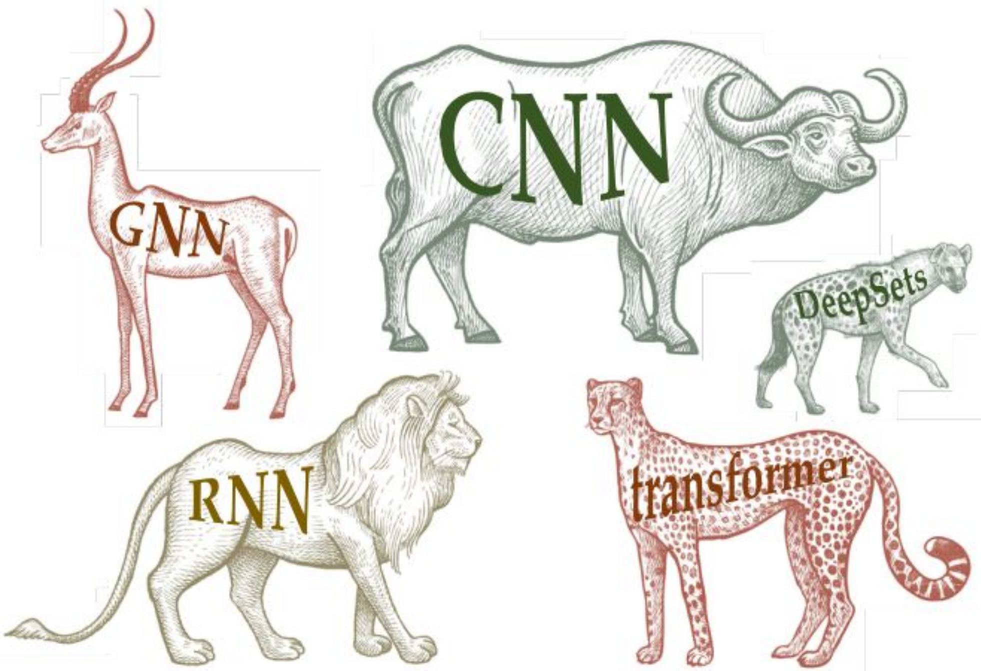


GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Programa de Erlangen

- F. Klein en 1872
- Estudiar variedades homogéneas por medio de sus propiedades que son invariantes bajo transformaciones.

Existe un “zoológico” de arquitecturas de redes neuronales sin principios unificados



GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Geometric Deep Learning

- Concepto introducido en 2017
- Autores: Michel M. Bronstein, Joan Bruna, Taco Cohen, Peter Veličković
- Desarrollaron este concepto para ampliar los métodos de aprendizaje profundo hacia datos con estructuras complejas, datos que no pueden ser representados en espacios euclidianos.

Desafío de los datos euclidianos

- Los modelos tradicionales de DL están diseñados para datos en espacios euclidianos.
- Muchas estructuras reales no siguen esta estructura regular.

Limitaciones de los enfoques tradicionales

- Las redes neuronales profundas en sus formas tradicionales no pueden aprovechar las propiedades geométricas de los datos no euclidianos.
- El aprendizaje de representaciones efectivas en estos dominios complejos requiere técnicas especializadas.

El uso de simetrías permite que los modelos de GDL aprendas representaciones más eficientes y generalizables para datos estructurados.

Desafíos que enfrentan los modelos de DL y cómo GDL puede ofrecer soluciones

1) Aprendiendo en altas dimensiones

- **Motivación:** Los datos no se encuentran en espacios de baja dimensión.
- **Problemas:**
    - En espacios de alta dimensión, los patrones y las relaciones entre los datos se vuelven más difusos.
    - Los modelos de aprendizaje son menos efectivos a medida que aumenta la dimensionalidad.

- **Motivación:** GDL aprovecha las *estructuras geométricas inherentes*, como la *simetría, la topología o la geometría de los datos*, para aprender representaciones más *compactas y significativas*.

2) Sesgo inductivo / sobreajuste

- **Motivación:**
  - Se refiere a la forma en que los modelos de aprendizaje automático generalizan a partir de ejemplos.
  - El modelo necesita un sesgo estructural o regularidad en los datos para poder aprender patrones efectivos sin sobreajuste
- **Problemas:**
  - Sobreajuste
- **Motivación:** En GDL, se utiliza el concepto de *regularidad geométrica* para introducir un sesgo inductivo que facilita el aprendizaje, se usan las simetrías y invariancias inherentes en los datos (rotaciones, traslaciones, etc.)

3) La maldición de la dimensionalidad

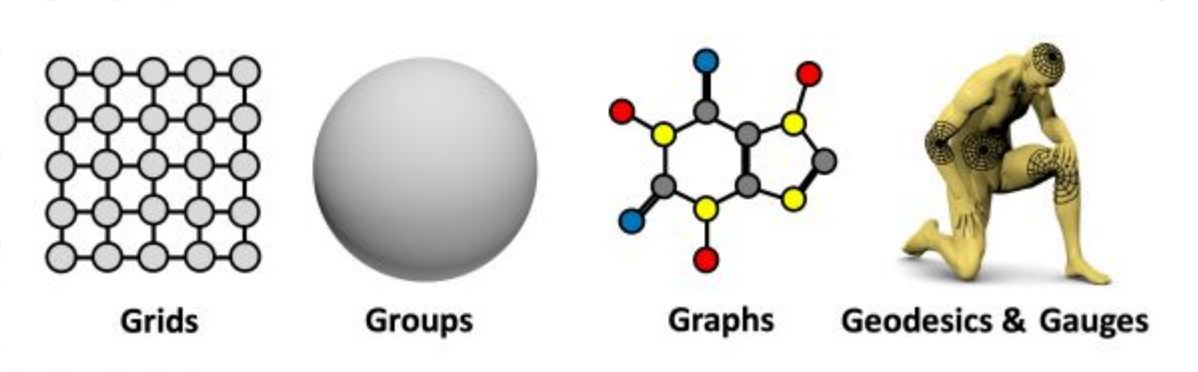
- **Motivación:** A medida que aumentan las dimensiones, los puntos de datos se dispersan cada vez más, lo que dificulta encontrar relaciones significativas.
- **Problemas:**
  - Varios algoritmos tradicionales sufren de esto porque la distancia entre los puntos se vuelve más uniforme a medida que aumenta la dimensión.
  - Los modelos tienen dificultades en identificar vecinos más cercanos o patrones significativos.

- **Motivación:** GDL combate la maldición de la dimensionalidad utilizando técnicas que explotan la *estructura geométrica de los datos*. No aprende en el espacio de alta dimensión original, encuentra representaciones más compactas e informativas de los datos.

Antecedentes geométricos

¿Dominio geométrico?

- En GDL los datos “viven” en un dominio Ω.
    - El dominio es un conjunto, con posiblemente una estructura adicional:

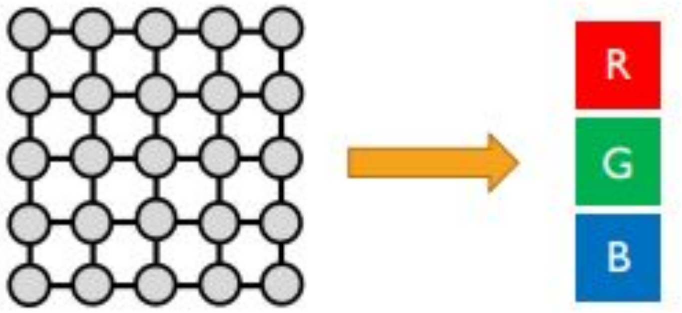


Mensaje clave: Las redes neuronales para procesar datos geométricos deben respetar la **estructura** del dominio.

GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Espacio de señales

- Una señal en Ω es una función
  $$x : \Omega \to \mathcal{C}$$
- El espacio de señales $\mathcal{C}$-valuadas
  $$\mathcal{X}(\Omega, \mathcal{C}) = \{x : \Omega \to \mathcal{C}\}$$
- $\mathcal{C}$ es un espacio vectorial cuyas dimensiones se llaman *canales*.



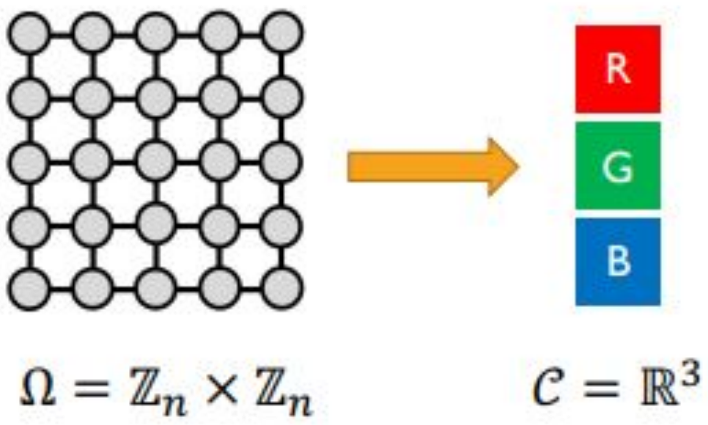
$\Omega = ?$ $\quad$ $\mathcal{C} = ?$

Ejemplo: imagen RGB de dimensión $n \times n$

GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Espacio de señales

- Una señal en Ω es una función
  $$x : \Omega \to \mathcal{C}$$
- El espacio de señales $\mathcal{C}$-valuadas
  $$\mathcal{X}(\Omega, \mathcal{C}) = \{x : \Omega \to \mathcal{C}\}$$
- $\mathcal{C}$ es un espacio vectorial cuyas dimensiones se llaman *canales*.



Ejemplo: imagen RGB de dimensión
$$n \times n$$

GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Estructura de espacio de “Hilbert”

- Las señales se pueden “sumar” y “multiplicar” por escalares:

$(\alpha x + \beta y)(u) = \alpha x(u) + \beta y(u)$ donde $\alpha, \beta \in \mathbb{R}$ y $u \in \Omega$

- El espacio de señales es un espacio vectorial.
- Se puede definir un producto interno: $\langle x, y \rangle = \int_\Omega \langle x(u), y(u) \rangle_\mathcal{C} \, \mathrm{d}\mu(u).$

GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Algunos conceptos

- Dominio Ω
- Señales (funciones) en el dominio
- Por ejemplo, consideremos un grid $n \times n$

$$\boldsymbol{\Omega} = \mathbb{Z}_n \times \mathbb{Z}_n$$

- Una imagen RGB x, entonces una señal:

$$x : \Omega \to \mathbb{R}^3$$

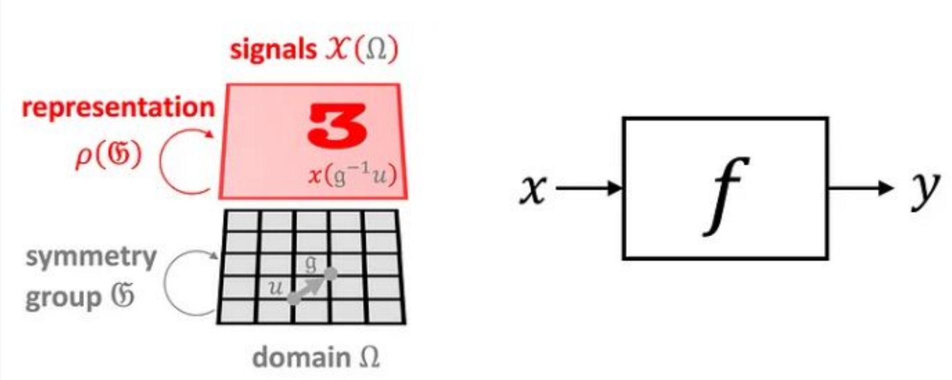
- Una función f como un Perceptrón de una capa que opera en un espacio de dimensión $3n^2$.

- El espacio de señales $\mathcal{C}$-valuadas en el dominio es un espacio de funciones con estructura de espacio vectorial:

$$\mathcal{X}(\Omega, \mathcal{C}) = \{x : \Omega \to \mathcal{C}\}$$

donde $\mathcal{C}$ es un espacio vectorial cuyas dimensiones se llaman *canales*.

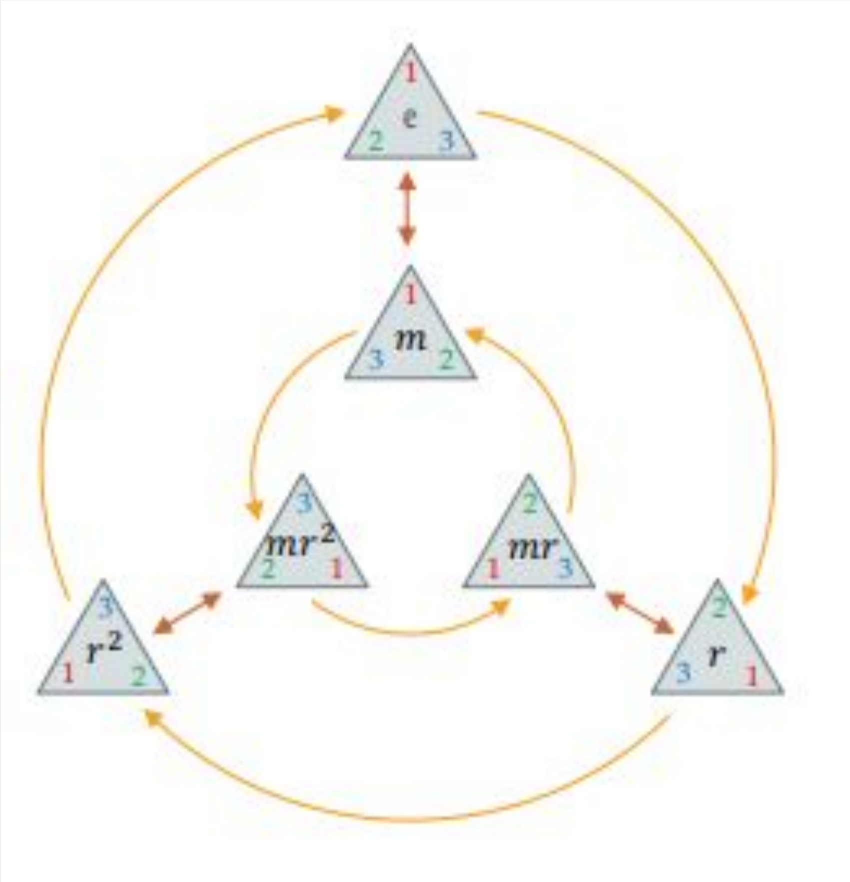
1)Simetrías, representaciones e invariancia



GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

¿Qué es una simetría?

Una simetría de un objeto, es una transformación del objeto que lo deja invariante.



GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

¿Ejemplos?

Ejemplos en ML: ____

¿Qué es una simetría? ¿Ejemplos?

Ejemplos en ML:

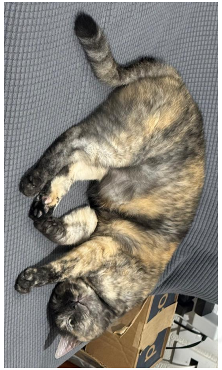
- Shifts
- predecir propiedades de moléculas independientemente de su orientación(rotaciones)

Grupo de simetrías

- Transformaciones de los datos que no cambian su significado o estructura.
- Transformación del espacio en sí mismo que preserva la naturaleza de la información.
- Ejemplos: traslaciones, rotaciones, reflexiones, permutaciones (en grafos).

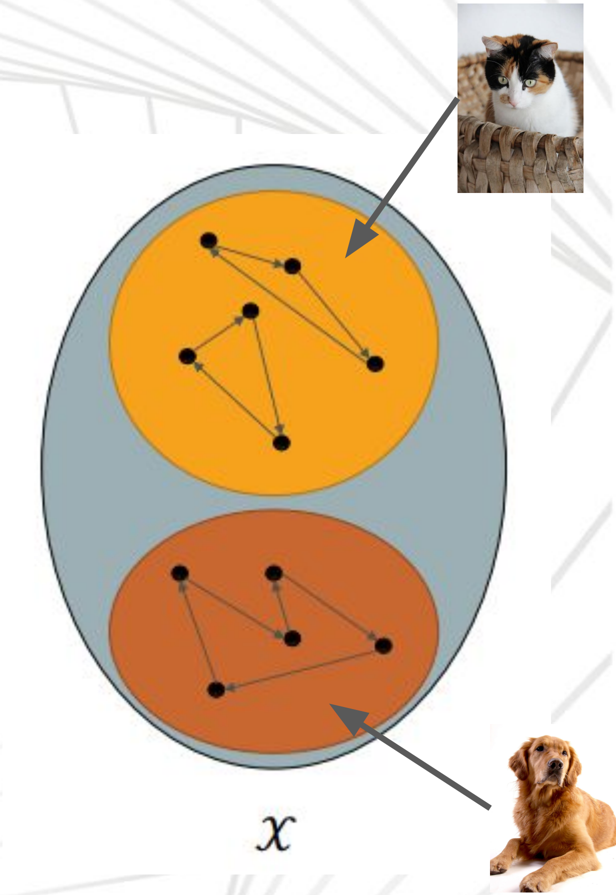
Simetrías de la función etiqueta

- En redes neuronales convolucionales CNN para imágenes, tenemos la traslación, mover un objeto en una imagen no cambia su esencia.
  - Sea X el espacio base y sea Y el espacio de etiquetas.
  - Sea $L : X \to Y$ la función de la etiqueta verdadera.

$L($  $) = L($  $) = gato$

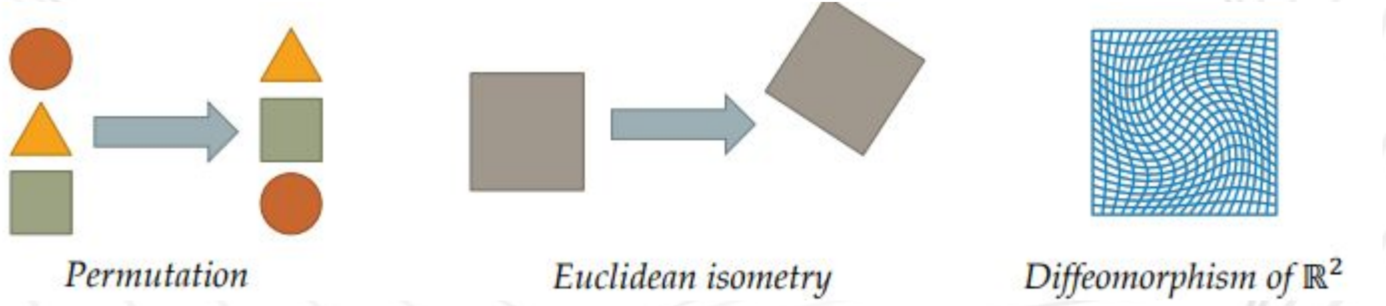
Aprender clases $\cong$ aprender simetrías

- Si sabemos todas las simetrías de la función L de etiquetas, entonces solo necesitamos saber una etiqueta por clase.
- Sin embargo, si el problema de aprendizaje es no trivial, no podemos esperar encontrar todas las simetrías del grupo.



Simetrías de dominios estructurados

- Una transformación $\mathfrak{g} : \Omega \to \Omega$ es una simetría si preserva la estructura del dominio.
- Ejemplos:
  - Permutaciones de los elementos de un conjunto
  - Isometrías euclidianas (rotaciones, traslaciones, reflexiones) preservan la distancia y los ángulos en el espacio euclidiano.
  - Un difeomorfismo preserva la estructura suave de una variedad.

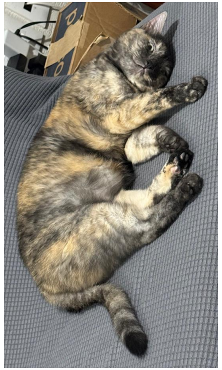
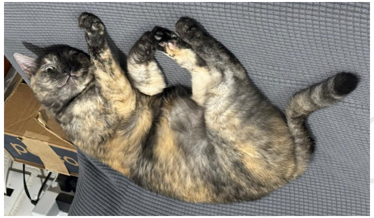


GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Grupo de simetrías

- La colección de todas las simetrías de un objeto dado se llama **grupo de simetrías.**
    - La transformación identidad es siempre una simetría.
    - Dadas dos simetrías, su composición es también una simetría.
    - Dada cualquier simetría, su inversa es también una simetría

Id 

 $\circ$  $= ?$

Grupos de simetrías, grupos abstractos y acciones de grupo.

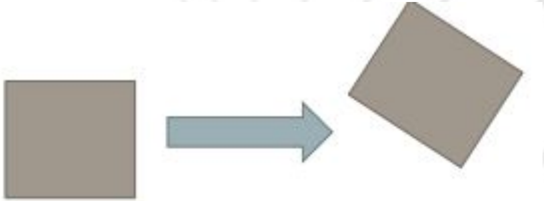
- Acciones de grupos:

$$
\begin{aligned}
G \times \Omega &\to \Omega \\
(g,\, u) &\to gu
\end{aligned}
$$

$$((\theta, t_x, t_y), (x, y)) \mapsto \begin{bmatrix} \cos\theta & \sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

tal que

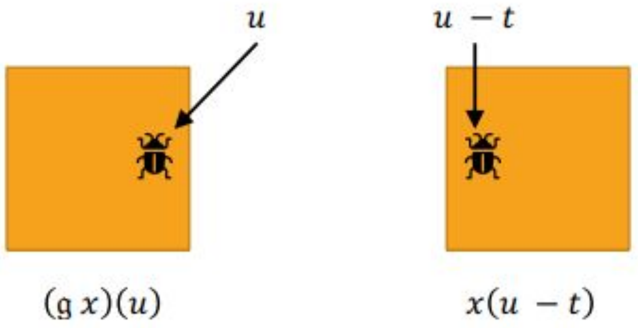
$$
\begin{aligned}
g\,(hu) &= (gh)u \\
eu &= u
\end{aligned}
$$

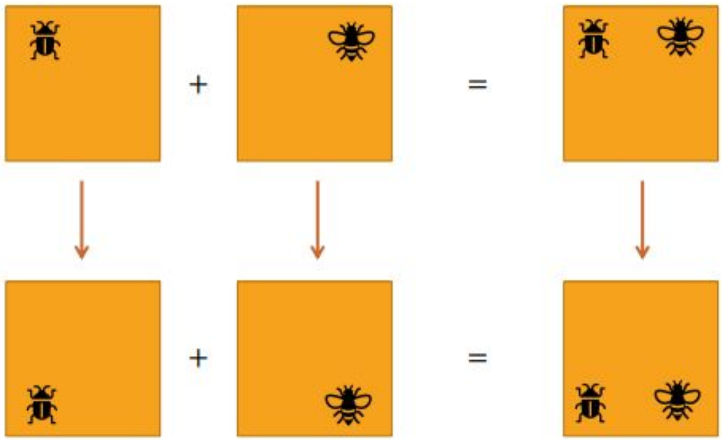


Simetrías actuando en señales

- Dada una acción, automáticamente obtenemos una acción en el espacio de señales: $(\mathfrak{g}\,x)(u) = x(\mathfrak{g}^{-1}u)$
- Ejemplo: traslación $\mathfrak{g} = (t_x, t_y)$

$$\mathfrak{g}\,(\alpha x + \beta y) = \alpha \mathfrak{g} x + \beta \mathfrak{g} y$$





GDL, Grids, Groups, Graphs, Geodesics and Gauges. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. Arxiv: 2104.13478v2

Representaciones del grupo

- Una representación de un grupo G es un mapeo $\rho: G \to \mathbb{R}^{n\times n}$ que asigna a cada $g \in G$ una matriz invertible $\rho(g)$ que satisface

$$\rho(gh) = \rho(g)\rho(h), \forall g, h \in G$$

- Ejemplo:
  - El grupo $G = (\mathbb{Z}, +)$
  - El dominio $\Omega = \mathbb{Z}_5 = \{1, 2, 3, 4, 5\}$
  - La acción de $g = n$ en $u \in \Omega : (n, u) \mapsto n + u \,(mod\, 5)$
  - La representación en $\mathcal{X}(\Omega)$

$$\rho(n) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}^n$$

$\rho(1)$  $=$

Representaciones del grupo

- Una representación de un grupo G es un mapeo $\rho: G \rightarrow \mathbb{R}^{n\times n}$ que asigna a cada $g \in G$ una matriz invertible $\rho(g)$ que satisface

$$\rho(gh) = \rho(g)\rho(h), \; \forall g, h \in G$$

- Ejemplo:
  - El grupo $G = (\mathbb{Z}, +)$
  - El dominio $\Omega = \mathbb{Z}_5 = \{1, 2, 3, 4, 5\}$
  - La acción de $g = n$ en $u \in \Omega : (n, u) \mapsto n + u \,(mod\, 5)$
  - La representación en $\mathcal{X}(\Omega)$

$$\rho(n) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}^n$$

Órbitas y relaciones de equivalencia

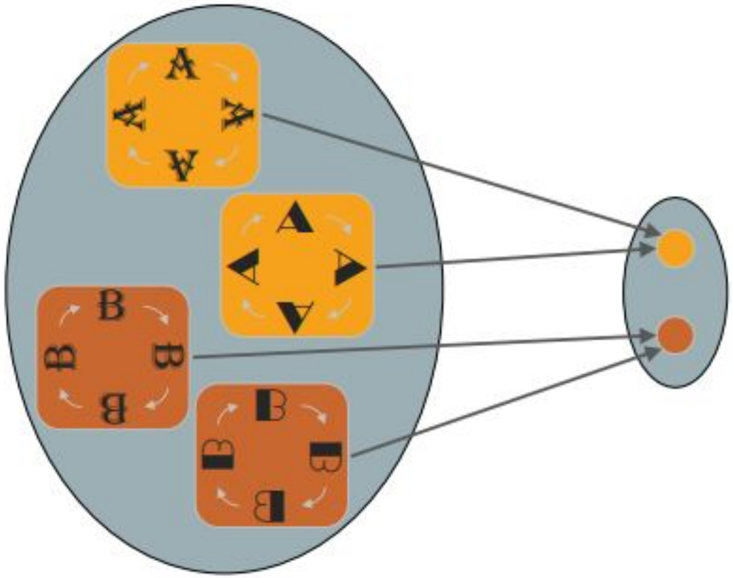
- $G-$equivalencia

$$x \sim_G y \iff g \in G : \; gx = y$$

y debe satisfacer los axiomas de una relación de equivalencia:

  - 1. ?
  - 2. ?
  - 3. ?

Órbita

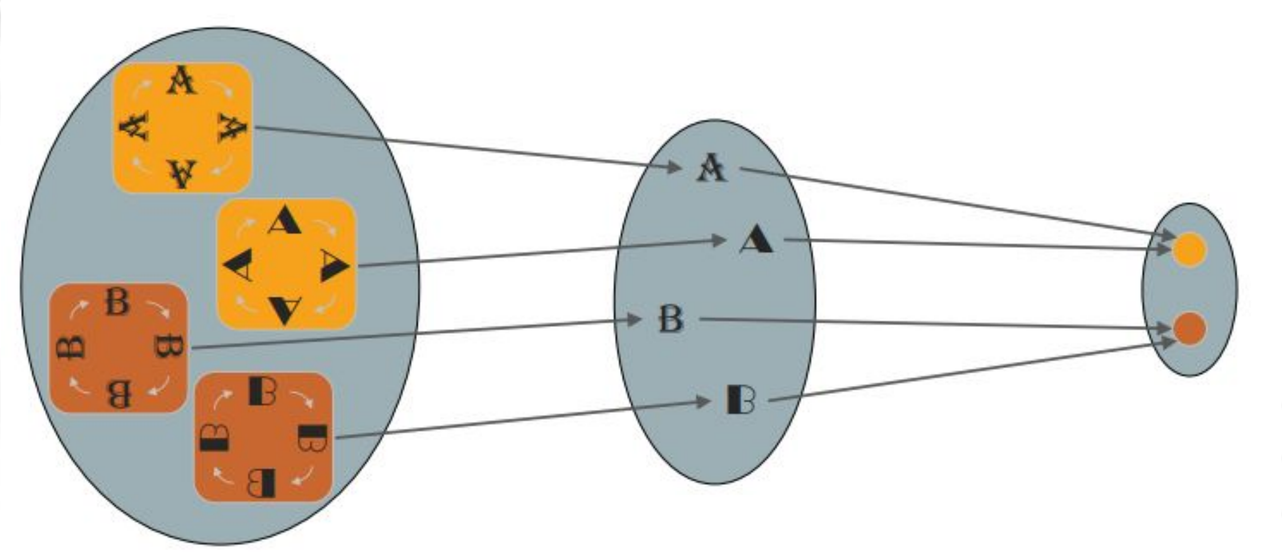


$$O_x = \{g\, x \mid x \in \mathcal{X}, g \in \mathfrak{G}\}$$

Representaciones invariantes

- Una función $f : \mathcal{X}(\Omega) \rightarrow \mathcal{X}(\Omega)$ es $G-$invariante si

$$f(\rho(g)x) = f(x)\, \forall g \in G, x \in \mathcal{X}(\Omega)$$



- Ejemplo: shift-invariance
- MLP no tienen esta propiedad.
- CNNs son shift-equivarint

El problema con la invarianza en DL

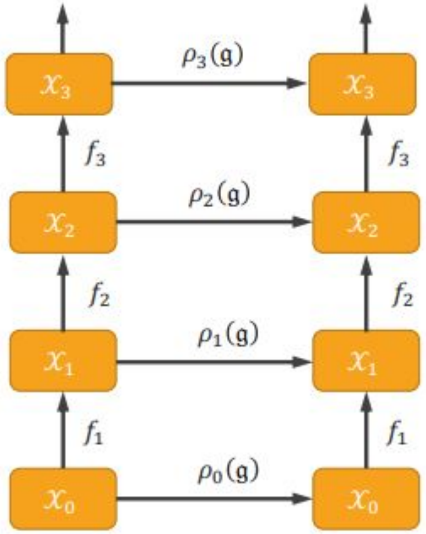
- Para reconocer objetos completos, primero necesitamos reconocer sus partes, por esta razón las redes neuronales deben ser profundas.
- Si hacemos representaciones invariantes intermedias, perdemos información

- Las posiciones relativas de las partes de los objetos contienen información crítica.

Redes equivariantes

- Una función $f : \mathcal{X}(\Omega) \to \mathcal{X}(\Omega)$ es $G-$equivariante si

$$f(\rho(g)x) = \rho(g)f(x)\,\forall g \in G$$

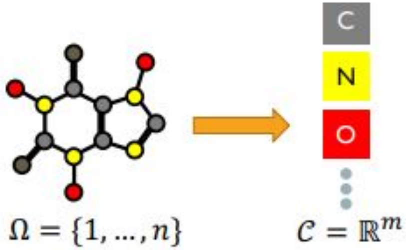


- Espacios vectoriales de características X_i
- Funciones entre las capas f_i
- Un grupo de simetrías G
- Representaciones del grupo ρ_i de G para cada X_i

Equivarianza: $f_i \circ \rho_{i-1}(g) = \rho_i(g) \circ f_i$

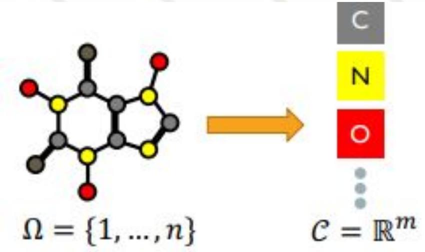
Ejemplo:

- Supongamos que con un modelo de GDL queremos clasificar moléculas.
- Cada molécula se representa como un grafo en el que los átomos son nodos y los enlaces son aristas.



1. **Simetrías:**
2. **Representación:**
3. **Invariancia:**

Ejemplo

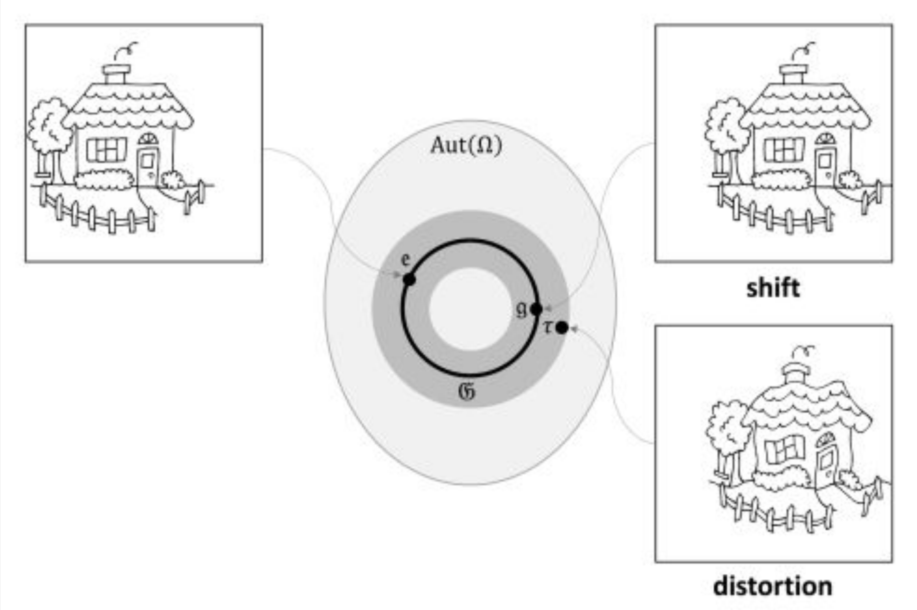


- Supongamos que con un modelo de GDL queremos clasificar moléculas.
- Cada molécula se representa como un grafo en el que los átomos son nodos y los enlaces son aristas.

1. **Simetrías:** La molécula tiene simetría bajo rotaciones y traslaciones en el espacio tridimensional, así como simetría bajo la permutación de los átomos idénticos.
2. **Representación:** Usamos una estructura de grafo para capturar las relaciones de enlace entre átomos, y representamos las simetrías de permutación en las operaciones de la red neuronal.
3. **Invariancia:** El modelo debe ser invariante a la rotación de la molécula y al orden en que los átomos están etiquetados. Esto asegura que siempre clasificará la molécula de la misma forma, independientemente de su orientación o etiquetado.

2) Isomorfismos y automorfismos

2) Estabilidad ante deformaciones



No estamos interesados en grandes conjuntos de transformaciones donde la invarianza global es exacta es reemplazada por una invarianza local inexacta

Estabilidad de señales de deformaciones

- Se refiere a la capacidad de reconocer patrones y relaciones geométricas en los datos a pesar de deformaciones locales o globales.
- Deformaciones pequeñas de señales no deberían cambiar el output.
- Cuantificar qué tan **<u>"lejos"</u>** un difeomorfismo dado $\tau \in \mathrm{Diff}(\Omega)$ está de un grupo de simetrías $G \subset \mathrm{Diff}(\Omega)$.
- Estabilidad ante deformaciones:

$$||f(\rho(\tau)x) - f(x)|| \leq Cc(\tau)\,\|x\|,\ \forall x \in \mathcal{X}(\Omega)$$

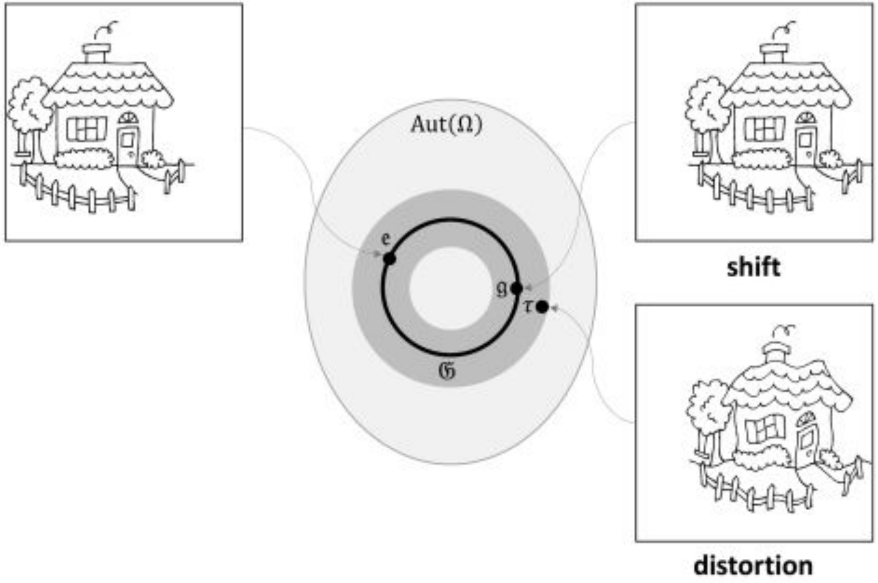
- C es una constante independiente de la señal x
- $c(\tau)$ es una medida de complejidad, tal que $c(\tau) = 0\ si\ \tau \in G$
- $\rho(\tau)x(u) = x(\tau^{-1}u)$

¿Costo?

- Se debe definir un costo apropiado de la deformación.
- Por ejemplo: para imágenes definidas sobre el plano Euclidiano, una elección suele ser

$$c^2(\tau) := \int_\Omega ||\nabla\tau(u)||^2 \mathrm{d}u$$

la cual mide la elasticidad de τ



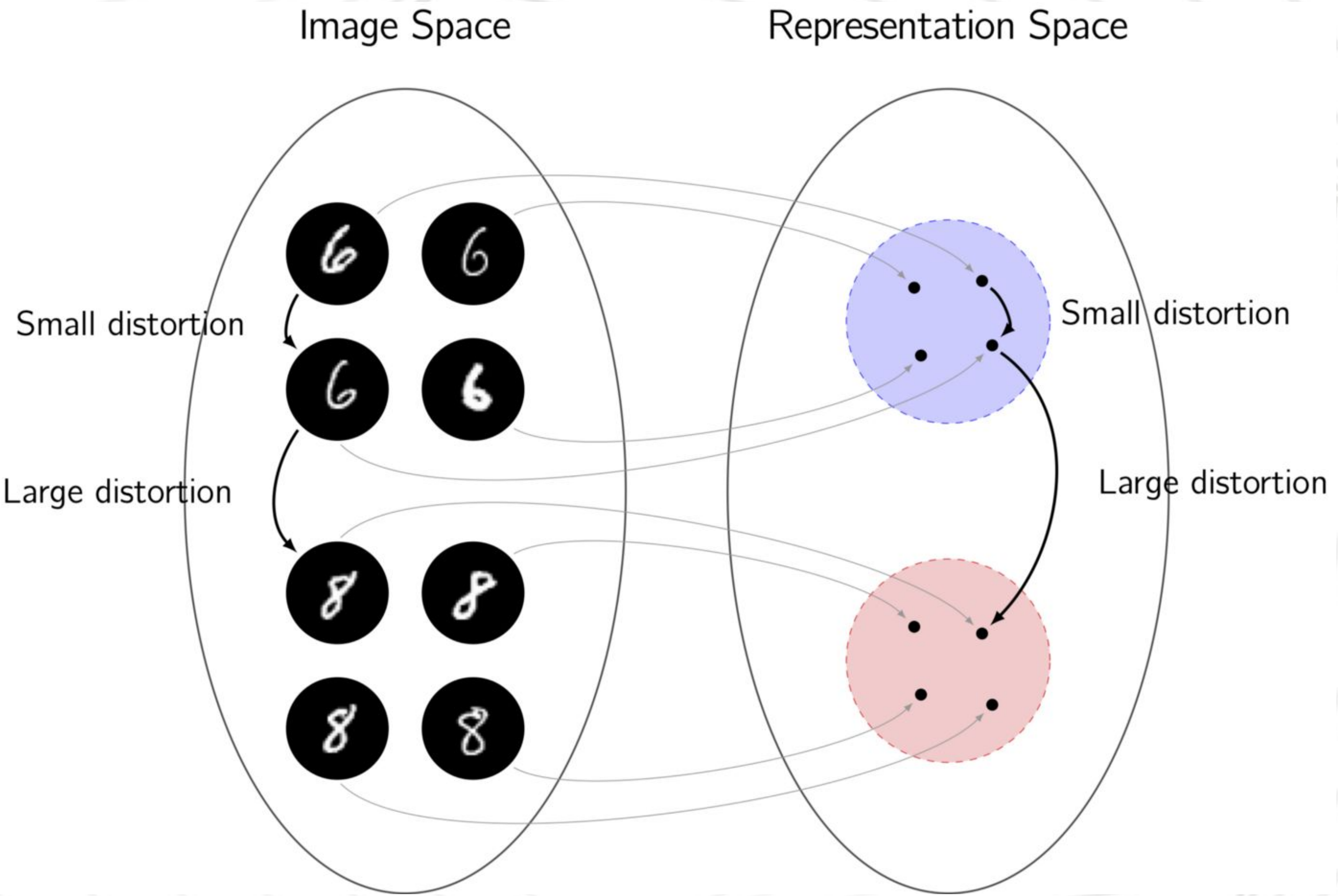
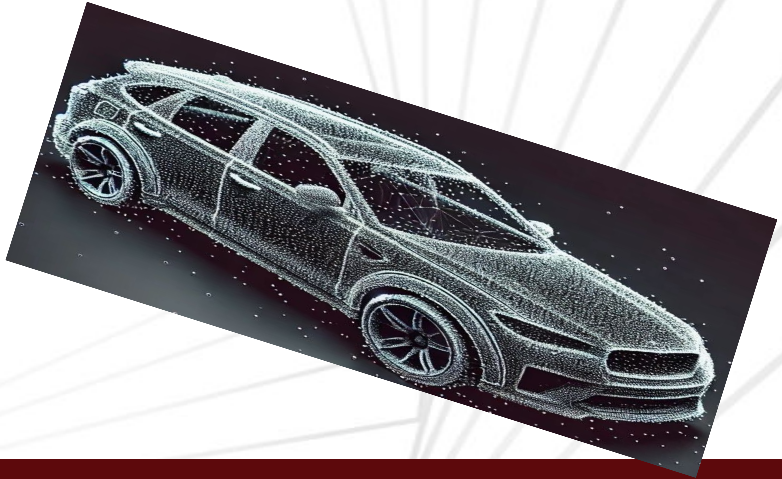


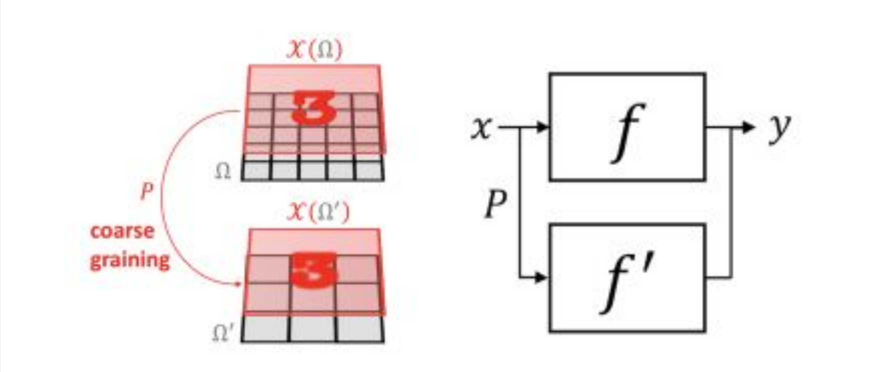
Image Space
Representation Space
Small distortion
Large distortion
Small distortion
Large distortion

Ejemplo

- Clasificar formas 3D de objetos.
- Modelo capaz de reconocer un coche
- Representamos un coche 3D por una nube de puntos.
- Si se rota o se deforma de alguna manera, el modelo debe seguir siendo capaz de reconocer que es un coche y clasificarlo correctamente.
- Se pueden emplear redes neuronales, que sean invariantes a transformaciones como rotaciones y escalas.
- Estas redes deben aprender características globales y locales de la nube de puntos.
- Deben ser capaces de mantener la estabilidad de la representación del objeto incluso frente a deformaciones.



3) Separación de escala

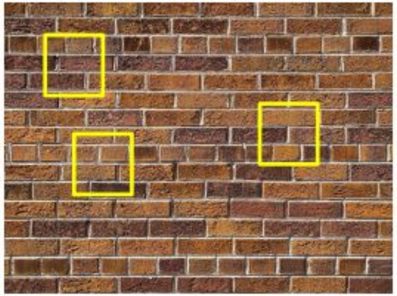


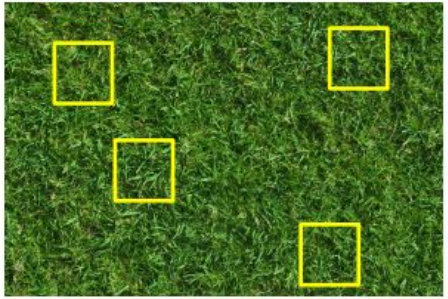
Capacidad de un modelo para manejar y aprender de información en diferentes escalas de la estructura geométrica.

Que el modelo pueda reconocer y procesar tanto detalles locales como patrones globales

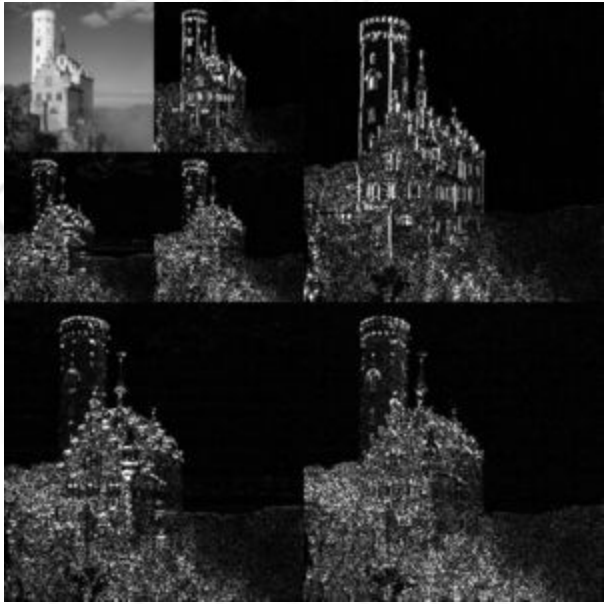
Ejemplos

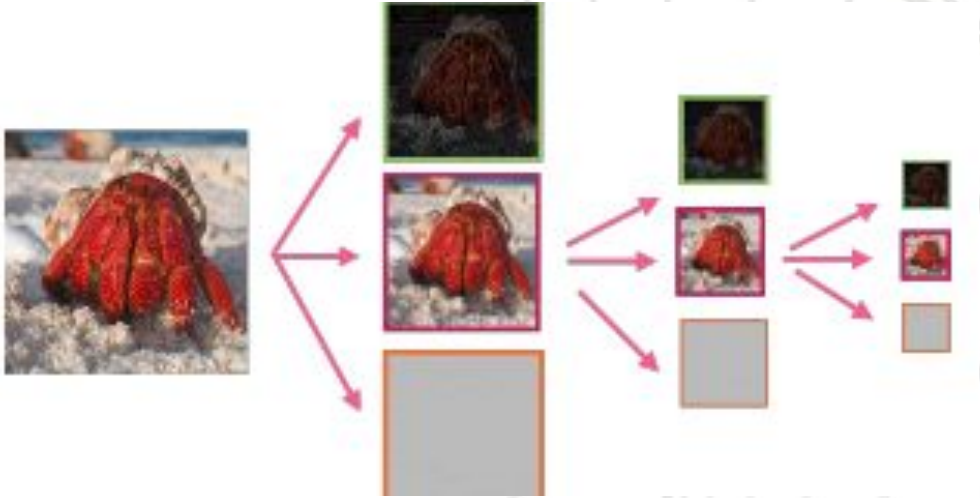
- Grafos: aprender características locales como la conectividad o las propiedades de los nodos vecinos y características globales como el flujo de información entre distintas comunidades dentro del grafo.
- Nubes de puntos 3D: las estructuras locales pueden incluir información de las posiciones de los puntos vecinos, mientras que las globales nos pueden capturar la forma de todo el objeto.



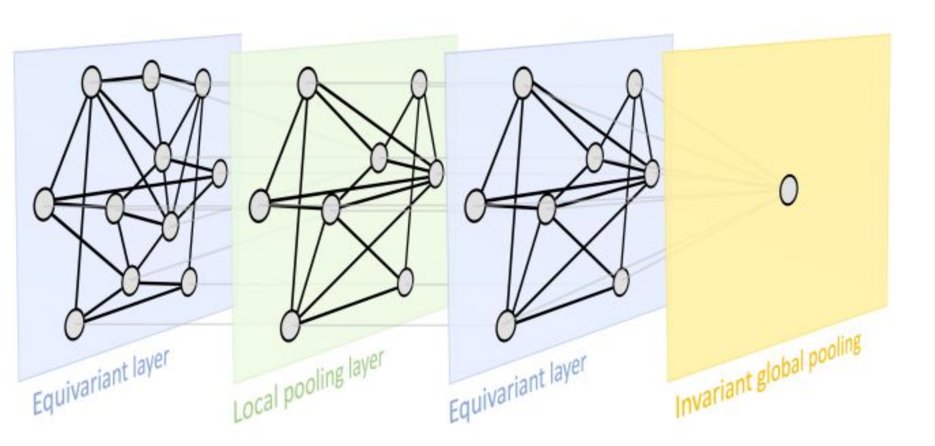


Representaciones multiescalas y análisis multiresolución





El “blueprint” de GDL



Bloques de construcción

- Combinar la simetría, la estabilidad geométrica y la separación de escalas nos pueden dar una representación estable de datos en alta dimensión.
- La geometría del dominio junto con un conocimiento del grupo de simetrías nos da los bloques de construcción:
  - Un mapeo local equivariante
  - Un mapeo global invariante
  - Un operador de simplificación (coarsening operator) <- reducir la resolución o complejidad de un estructura.

Blueprint del GDL

Sean Ω y Ω' dominios, G un grupo de simetría sobre Ω, y escribimos $\Omega' \subseteq \Omega$ si Ω' puede considerarse una versión compacta de Ω.

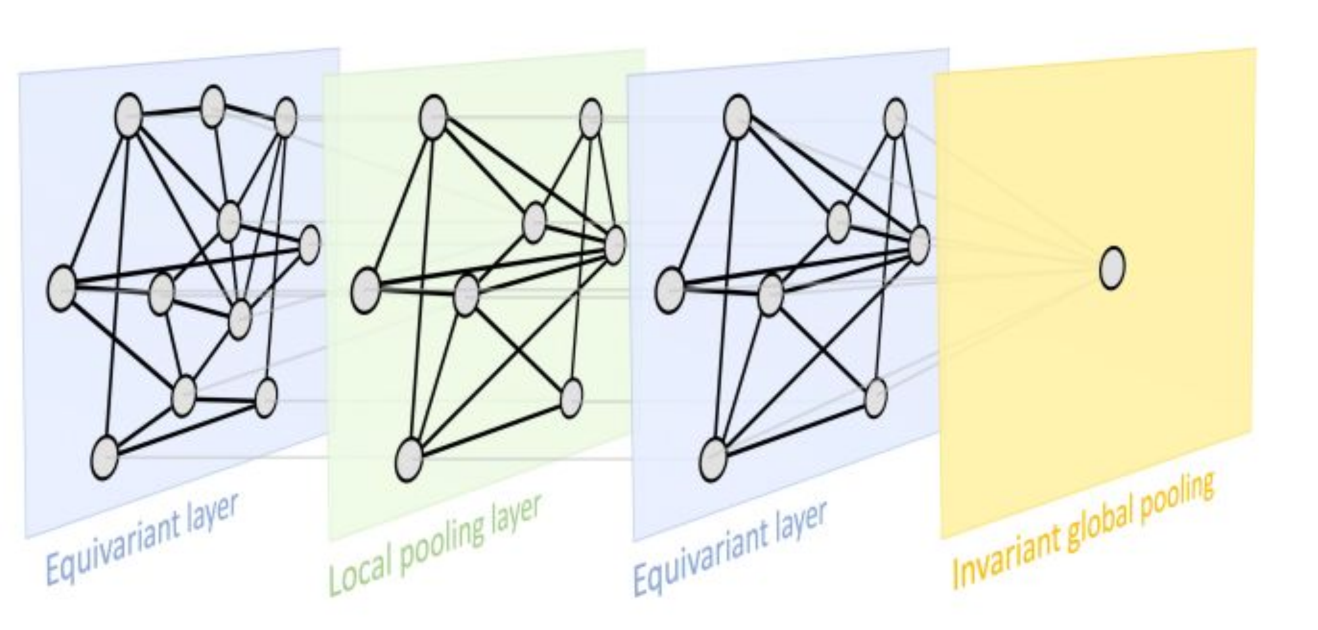
Definimos los siguientes bloques de construcción:

- **Capa lineal G-equivariante** $B : \mathcal{X}(\Omega, \mathcal{C}) \to \mathcal{X}(\Omega', \mathcal{C})$ satisfaciendo $B(g.x) = g.B(x)$ para todo $g \in G$ y $x \in \mathcal{X}(\Omega, \mathcal{C})$.
- **No linealidad** $\sigma : \mathcal{C} \to \mathcal{C}$ aplicada elemento a elemento como $(\sigma(x))_u = \sigma(x(u))$.
- **Agrupamiento local (coarsening)** $P : \mathcal{X}(\Omega, \mathcal{C}) \to \mathcal{X}(\Omega', \mathcal{C})$ tal que $\Omega' \subset \Omega$.
- **Capa G-invariante (agrupamiento global)** $A : \mathcal{X}(\Omega, \mathcal{C}) \to \mathcal{Y}$ que satisface $A(g.x) = A(x)$ para todo $g \in G$ y $x \in \mathcal{X}(\Omega, \mathcal{C})$.

Usando estos bloques es posible construir funciones G-invariantes $f : \mathcal{X}(\Omega, \mathcal{C}) \to \mathcal{Y}$ de la forma

$$f = A \circ \sigma_0 \circ B_J \circ P_{J-1} \circ \sigma_{J-1} \circ \ldots \circ P_1 \circ \sigma_1 \circ B_1$$

donde los bloques son seleccionados de manera que el espacio de salida de cada bloque coincida con el espacio de entrada del siguiente. Bloques diferentes pueden explotar distintas opciones de grupos de simetría G.

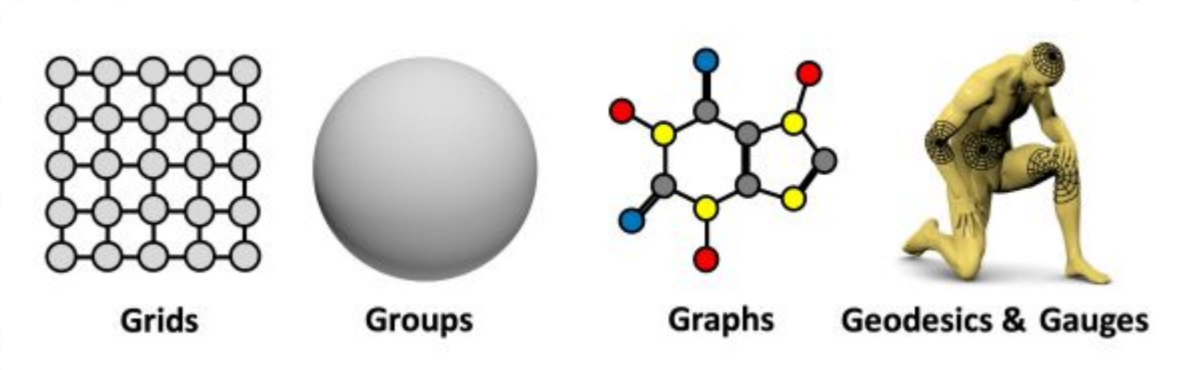


Un ejemplo de una arquitectura de una GNN que contiene una capa con una permutación equivariante (característica nodo a nodo), un local pooling y una permutación invariante global en la última capa.

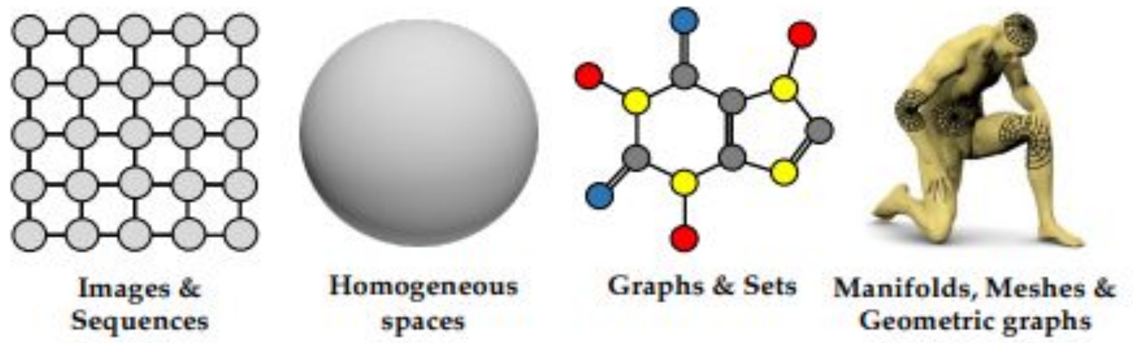
Ejemplos

Architecture	Domain Ω	Symmetry group $\mathfrak{G}$
CNN	Grid	Translation
Spherical CNN	Sphere / SO(3)	Rotation SO(3)
Intrinsic / Mesh CNN	Manifold	Isometry Iso(Ω) / Gauge symmetry SO(2)
GNN	Graph	Permutation Σ_n
Deep Sets	Set	Permutation Σ_n
Transformer	Complete Graph	Permutation Σ_n
LSTM	1D Grid	Time warping

Dominios geométricos: las 5Gs



Dominios geométricos: las 5Gs



Tarea:

- Ver la página https://geometricdeeplearning.com/
- Ver el video de la Lecture 1: https://www.youtube.com/watch?v=5c_-KX1sRDQ&feature=youtu.be