幾何学的な表現学習

福水健次
統計数理研究所

機械学習若手の会（YAML）2024
2024.9.22@ 浜名湖

自己紹介

1989. 京都大学理学部卒業（主として数学）

1989-1998. （株）リコー，研究開発本部
文字認識，音響信号処理，ニューラルネットの研究・開発

1996. 京都大学博士（理学）

1999-2000. 理化学研究所・脳科学総合研究センター・
甘利俊一研究室・研究員　ニューラルネットの統計的研究

2000-現在. 統計数理研究所
機械学習，数理統計の研究

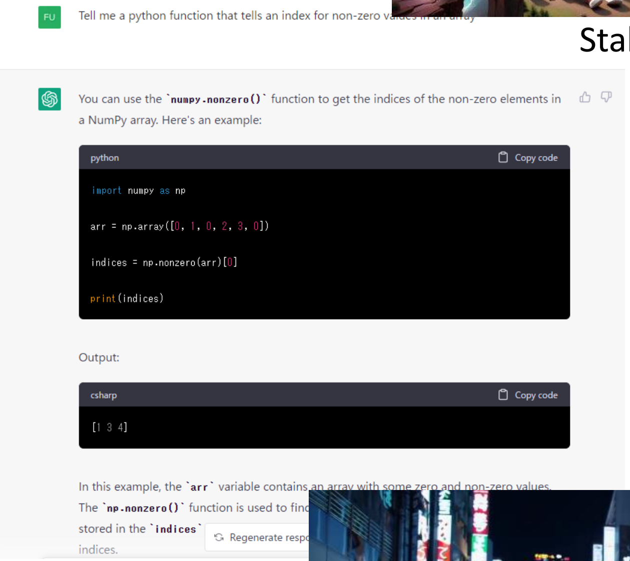
2018ー (株)Preferred Networks 技術顧問

2020 – JST CREST
「数理知能表現による深層構造学習モデルの革新」

機械学習の現在

- 深層学習を中心とする人工知能技術の革新的発展
  - 画像, 自然言語処理, 音声, 自動運転, ロボット, …
- 深層生成モデル
  - 画像生成
    Stable Diffusion, Midjourney, …
  - 大規模言語モデル
    ChatGPT, Gemini, LlaMa, o1, …
  - 動画生成
    Sora


Stability.AI, Stable Diffusion XL

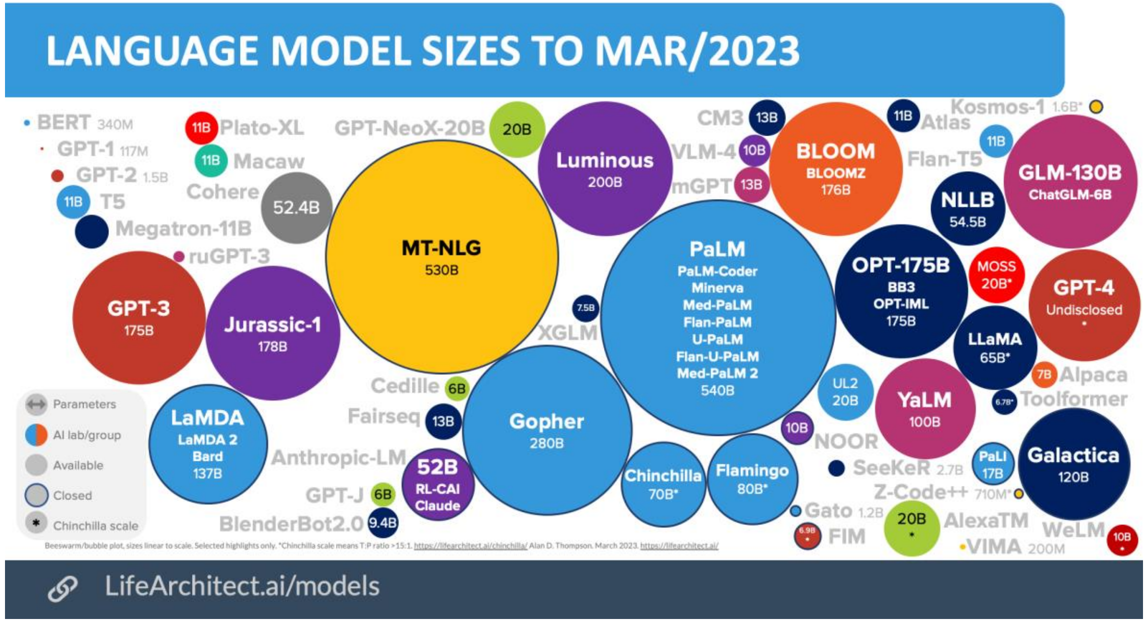



OpenAI, Sora

人工知能 vs 数学的知

- 深層学習による人工知能： 経験（データ）的知識獲得
  - Large Data / Large Model
    - **データからの知識獲得**
    - **表現学習：Implicitな情報表現**

      大量データにより
      有効な表現を自動獲得



Dr Alan D. Thompson, LifeArchitect.ai (Mar/2023).

- 数学的知： 演繹的知

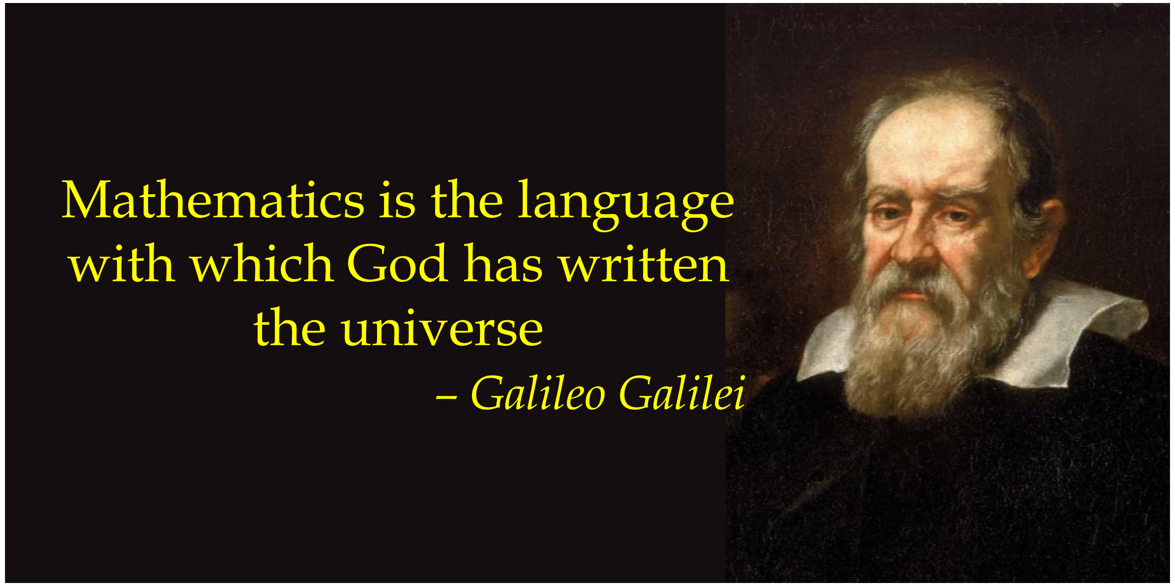
科学的知識は，伝統的には
「数学」により記述されてきた

e.g.
微分方程式

$$\rho \frac{D\boldsymbol{u}}{dt} = -\nabla p + \mu \nabla^2 \boldsymbol{u} + \frac{1}{3}\mu\nabla(\nabla \cdot \boldsymbol{u}) + \rho \boldsymbol{g}$$

代数
構造を抽象化する

$$\begin{tikzcd}
& A \arrow[r, "f"] \arrow[d, "a"] & B \arrow[r, "g"] \arrow[d, "b"] & C \arrow[r] \arrow[d, "c"] & 0 \\
0 \arrow[r] & A' \arrow[r, "f'"] & B' \arrow[r, "g'"] & C' &
\end{tikzcd}$$



- 人工知能＋数学的知

数学的知により，経験的知である人工知能技術をどのように変革できるか？
代数構造，幾何学，微分方程式，etc

講演の概要

1. イントロ
2. 対称性と群の作用
3. 機械学習と群作用
4. 群の表現
5. Neural Fourier Transform
6. まとめ

対称性と群の作用

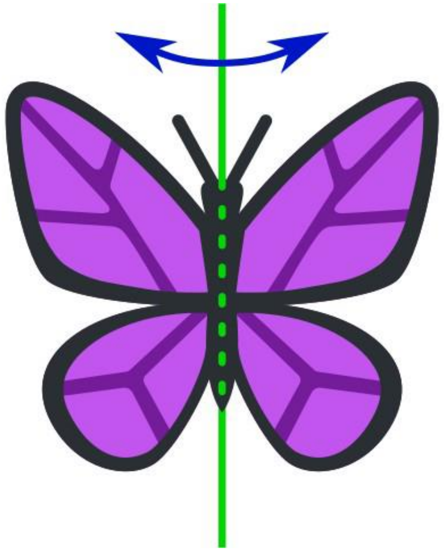
対称性

デザイン

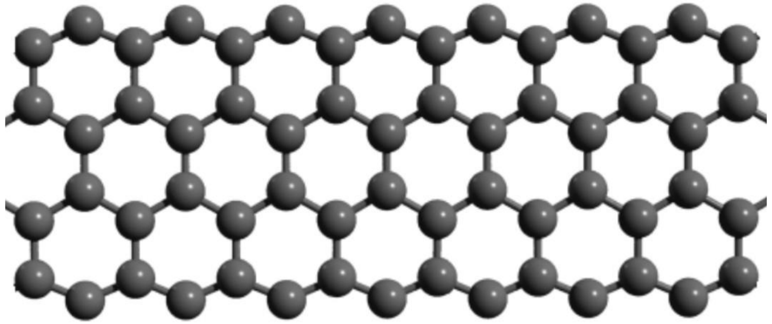
葵の御紋

Cloisters, University of Glasgow

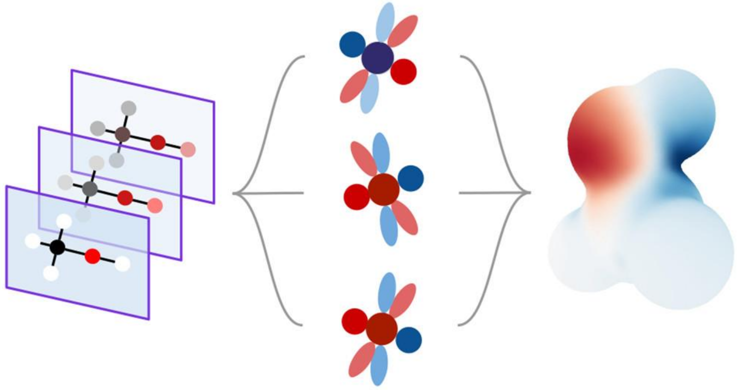
自然界



形態の対称性



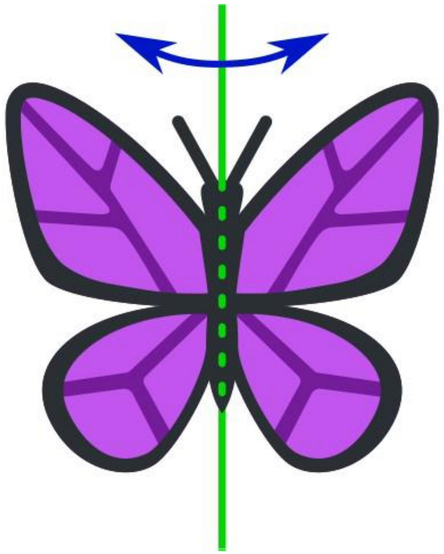
結晶格子



[Thürlemann et al. *J. Chem. Theory Comput.* 2022]

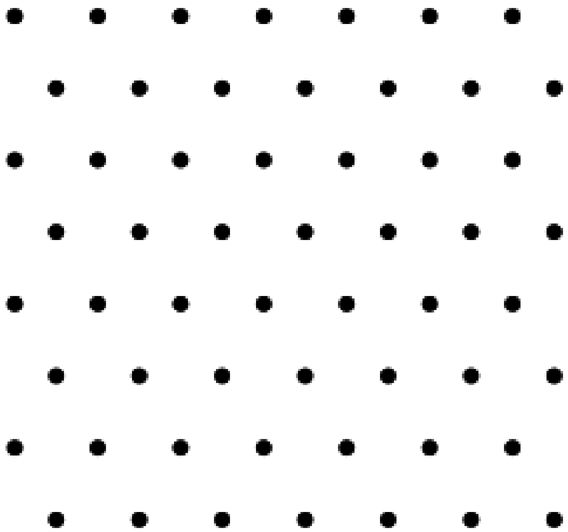
原子配置/ポテンシャル

対称性をとらえる変換



鏡映

回転



平行移動

群と群の作用

- 群：　対称性とそれをとらえる変換を数学的に定義したもの

定義　群（group）

空でない集合 G に演算が定義されている(積で表記) とき，　G が 群 であるとは,
(1) (結合律)　$(ab)c = a(bc)$　for any $a, b, c \in G$.
(2) (単位元の存在) $e \in G$ があって，$ae = ea = a$ for any $a \in G$.
(3) (逆元の存在)　任意の$a \in G$, に対し，$a^{-1} \in G$があって$aa^{-1} = a^{-1}a = e$.

例1）$(\mathbb{Z}^n, +)$, $(\mathbb{R}^n, +)$ 加法に関して可換な群（Abel群）．　$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$.

例2）$GL(\mathbb{R}^m)$ 一般線形群（可逆な$m \times m$行列全体）積に関して群.
非可換 $AB \neq BA$.
部分群　$O(m)$ 回転群

- 群の作用

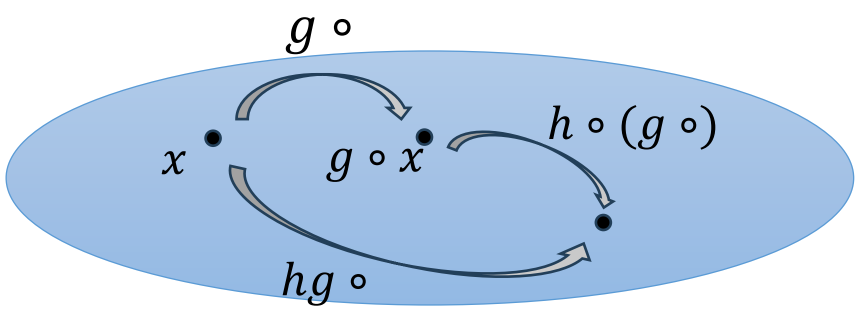
  定義

  G: 群, X: 集合.

  G の X への群作用 とは，各 $g \in G$ に対して X の変換 $g \circ$ が定まって

  i)　$e \circ x = x$　（$e \in G$は単位元）

  ii)　$(hg) \circ x = h \circ (g \circ x)$



- 典型的な例：

  1）$GL(\mathbb{R}^m)$ の $\mathbb{R}^m$ への標準的な作用　$A \circ x := Ax$　（行列とベクトルの積）
  　$O(m)$: 回転群

  2）シフト： $G = (\mathbb{R}^m, +)$ の$\mathbb{R}^m$自身への作用　$a \circ x := x - a.$

  3）$\mathcal{F} = \{f: G \to \mathbb{R}^m\}$（群上の写像）

  $$L_g(f) := [h \mapsto f(g^{-1}h)]$$

  L_g は G の$\mathcal{F}$への作用を定める.

不変性，同変性

定義.

群 G が集合 X と Y に作用． 写像 $\varphi: X \to Y$.

- φ がGの作用に対して不変（invariant）とは,

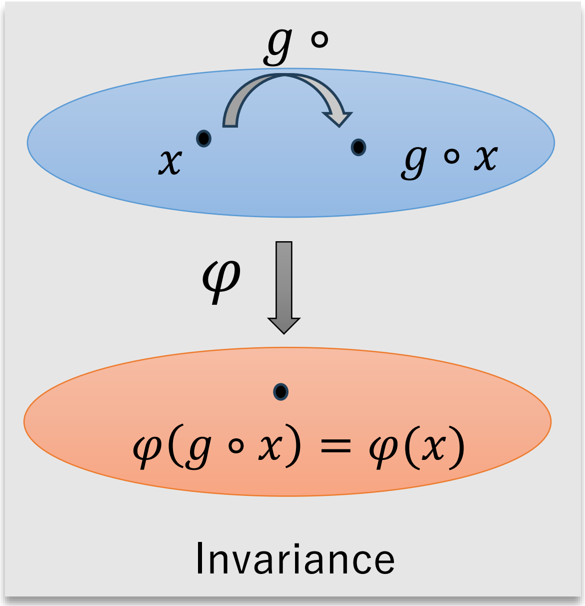
$$\varphi(g \circ x) = \varphi(x)$$

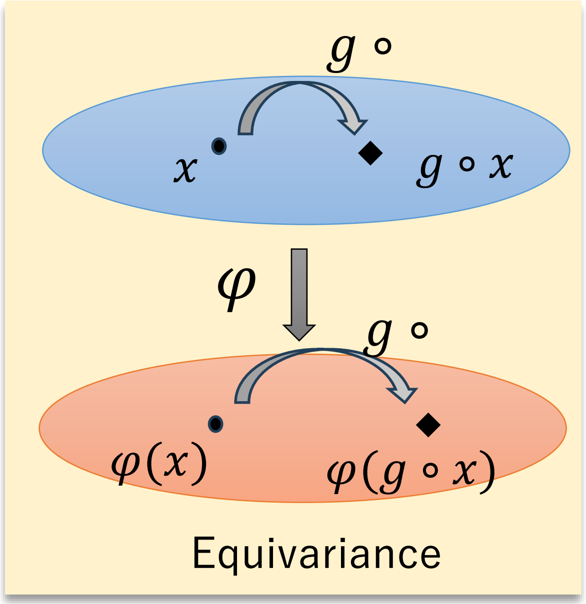
が任意の $g \in G, x \in X$ に対して成り立つこと.

- φがGの作用に対して同変(equivariant)とは

$$\varphi(g \circ x) = g \circ \varphi(x)$$

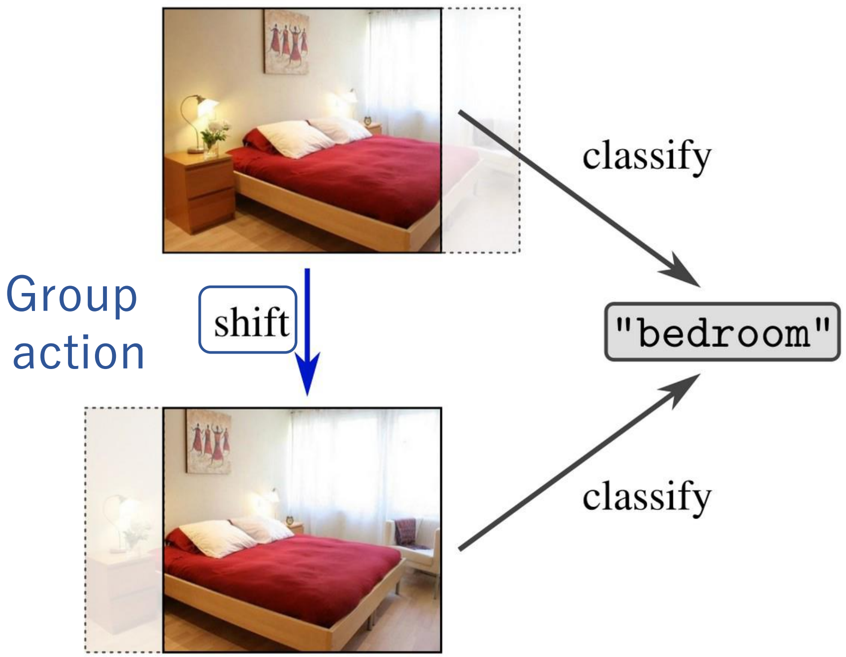
が任意の $g \in G, x \in X$ に対して成り立つこと.



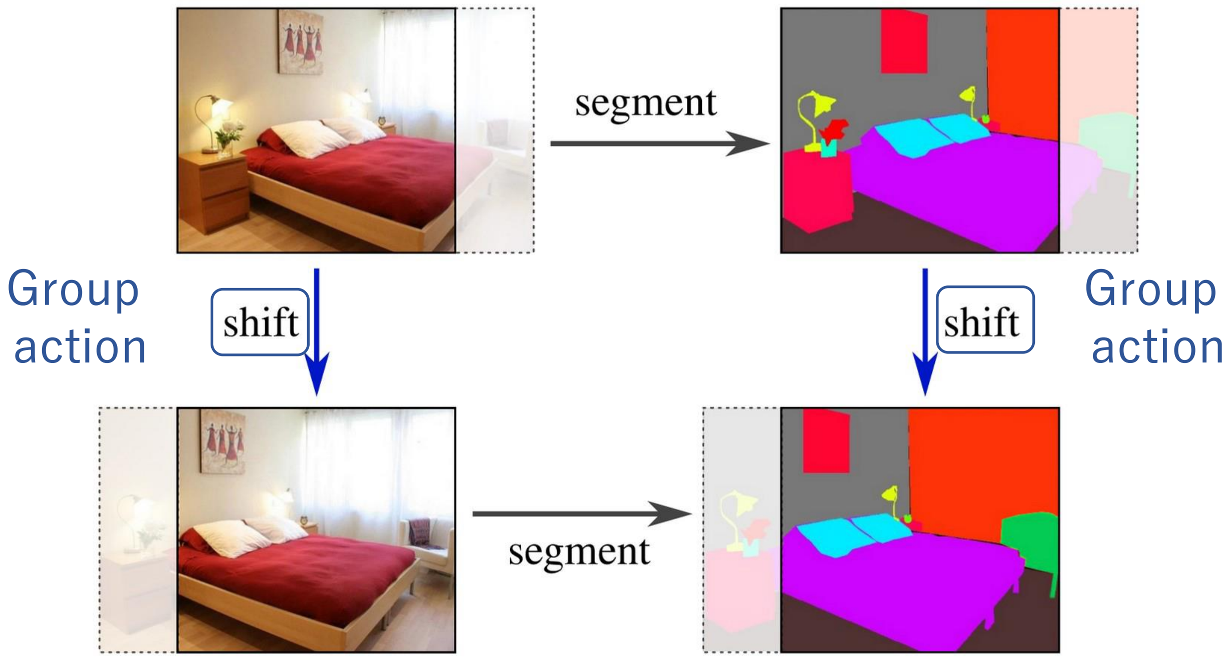


機械学習における群作用

不変性
object classification



同変性
segmentation



From “Groups, Representations & Equivariant maps” by Maurice Weiler (University of Amsterdam)

機械学習と群作用

不変性/同変性を用いた深層学習

不変性/同変性： 多くの有用なモデルで利用

- **Architecture:** モデル構造により同変性実現
  - Convolutional Neural Network (CNN):

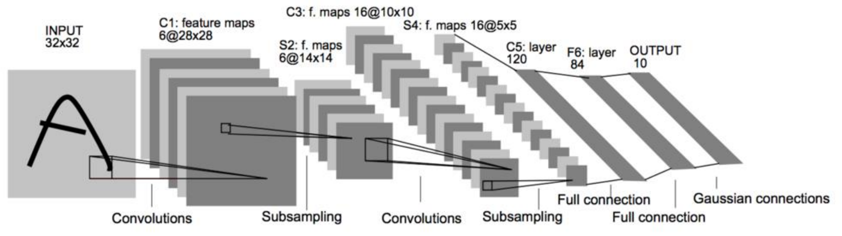
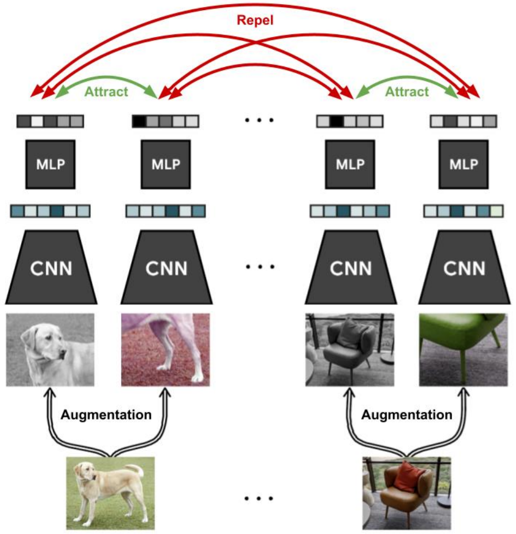


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

LeCun 1998

- **Data augmentation：** データ拡張に利用
  - データを変換により拡張して学習する.
  - 不変な識別機の学習
  - 「自己教師あり学習」．変換したものは近い表現



SimCLR

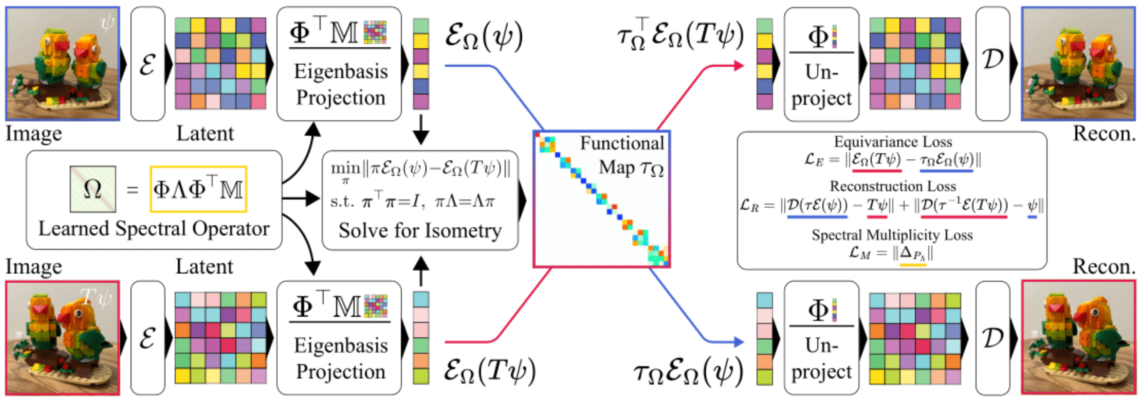
これらのアプローチでは群 G はknown

• 同変性に基づく表現学習

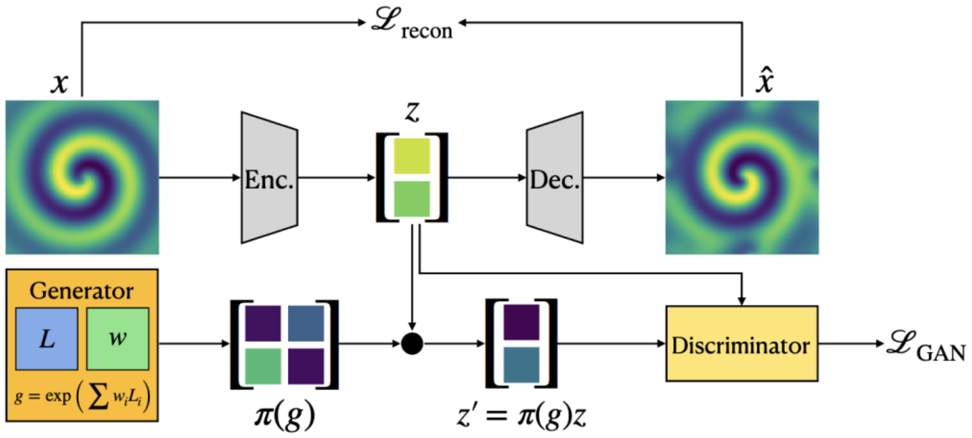
同変になるように学習を行うことによってよい表現を得る.

今日の本題

- Meta-sequential prediction (Miyato, Koyama, Fukumizu NeurIPS 2022)
- Neural Fourier Transform (Koyama, Fukumizu, Hayashi, Miyato ICLR 2024)
- Neural Isometries. (Mitchel et al arXiv 2024)
- Latent Space Symmetry Discovery. (Yang et al ICML 2024)



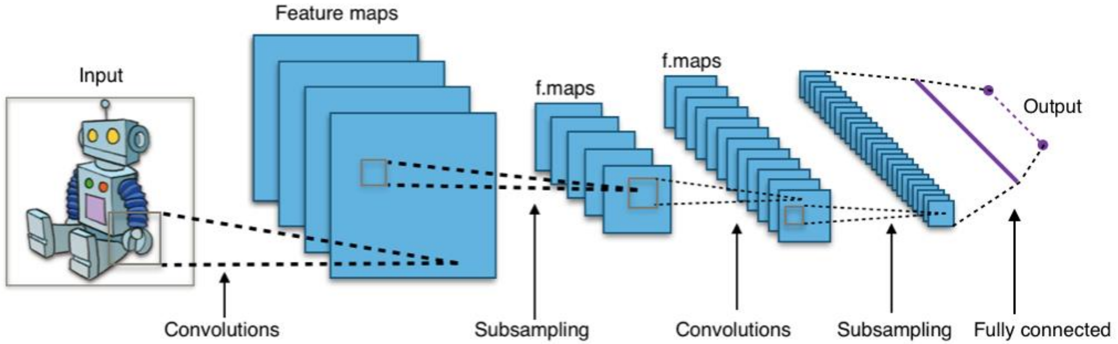
Neural Isometry

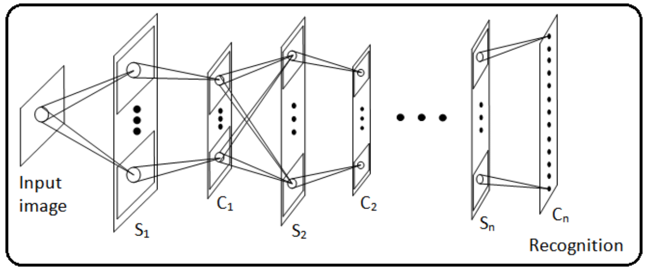


Yang et al ICML 2024

Convolutional Neural Networks

- CNN (Wei Zhang et al 1988; Yann LeCun 1989)
  - 平行移動同変性を実現するNNアーキテクチャ. 画像入力を想定
  - 畳み込み層+ Pooling層
    Backpropで学習.
  - 生物の初期視覚野の神経回路網
    モデルに触発されている.
  - オリジン: Neocognitron (Fukushima 1980)
    Backprop は使われていなかった.





Kunihiko Fukushima 1979

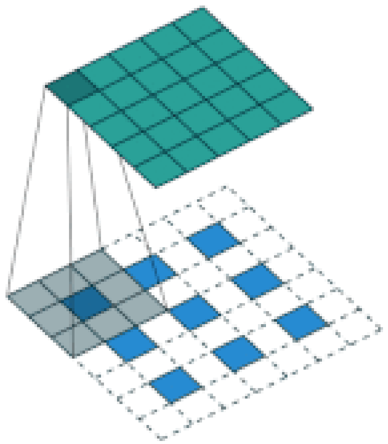
画素$[i,j]$の値 $f[i,j]$

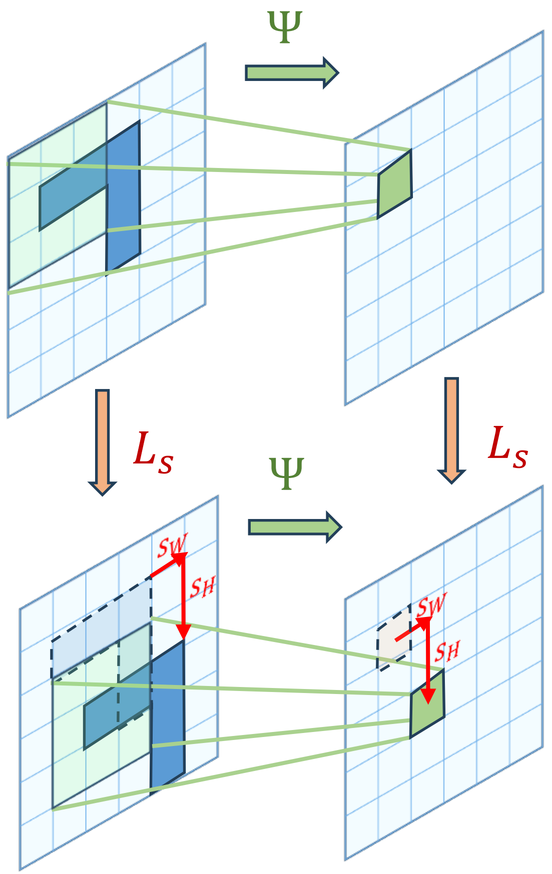
- 畳み込み層
  - 2D グレイスケール画像 $(W \times H)$

    $\psi_{[a,b]}$ ：空間フィルタ(3 × 3 or 5 × 5).

$$h^{Out}_{[i,j]} = \sum_{i-a\in\{0,\pm 1\}} \sum_{j-b\in\{0,\pm 1\}} \psi_{[i-a,j-b]} f^{In}_{[a,b]}$$

$$f^{Out}_{[i,j]} = \phi\left(h^{Out}_{[i,j]} + \theta\right) \qquad \phi\text{: 活性化関数}$$

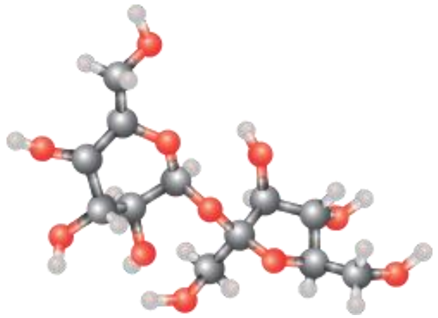




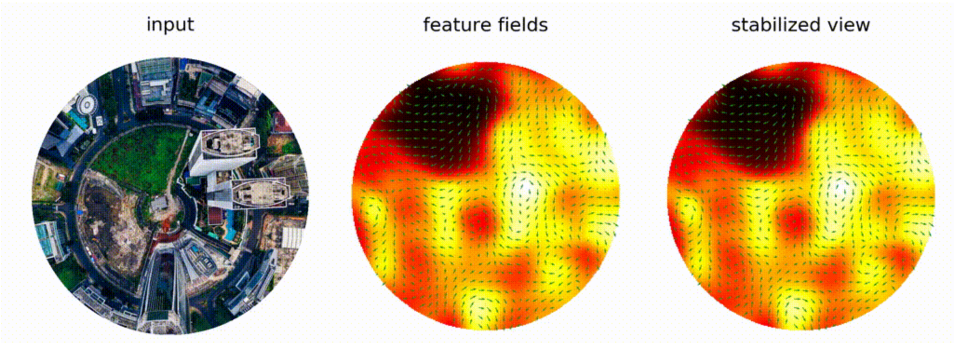
- 畳み込み層，Pooling層は同変写像

Group equivariant Convolutional Networks
(Cohen & Welling ICML 2016)

- G-CNN: さまざまな群Gの作用に対して同変なCNN
  - CNN
    - 画像：２次元格子からの写像.
      $$f: \mathbb{Z}^2(\text{画素}) \to \mathbb{R}^3(\text{RGB}).$$
    - $\mathbb{Z}^2$のシフトに対してCNNは同変.
      $$\Psi_{\text{Conv}}\left(f(\cdot - s)\right) = (\Psi_{\text{Conv}} f)(\cdot - s).$$



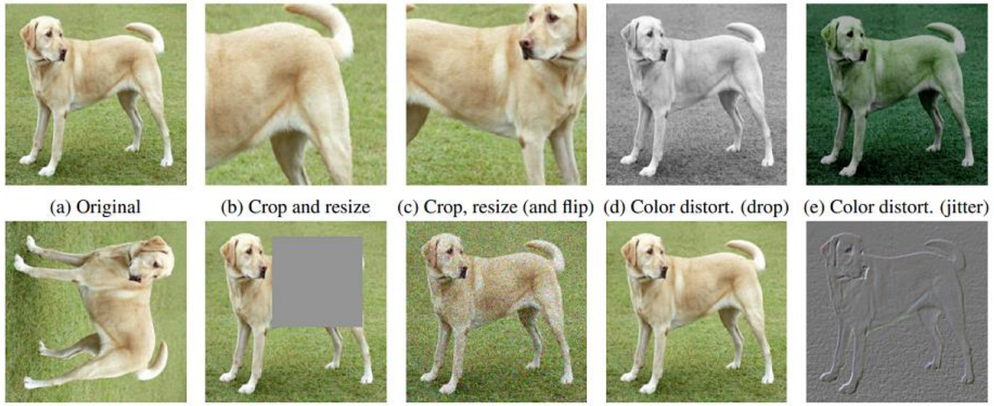
  - G-CNN
    - 入力信号：一般の群Gからの写像.
      $$f: G \to \mathbb{R}^K.$$
    - G から決まるシフト作用に関して同変.
      $$\Psi_{\text{Conv}}\left(L_g f\right) = L_g(\Psi_{\text{Conv}} f).$$

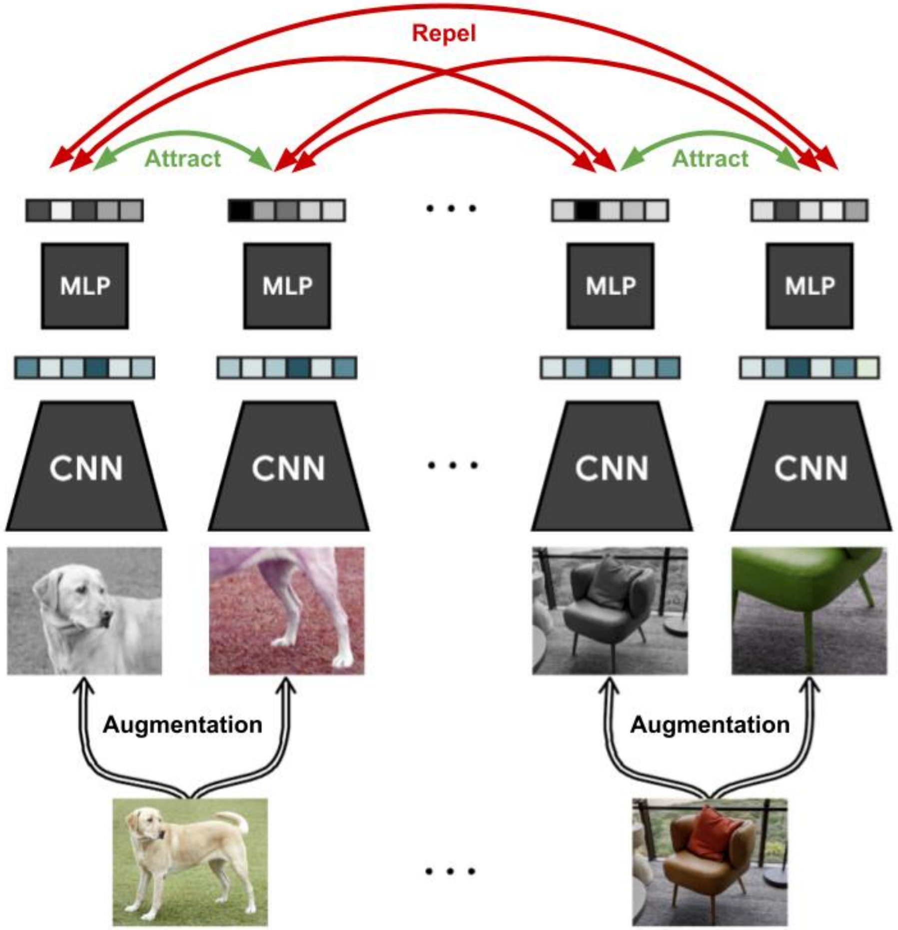


https://github.com/QUVA-Lab/e2cnn

自己教師あり学習：データ拡張

- 教師ラベルのない画像データからの特徴学習
- データ拡張：色変化，切り抜き，回転など
  （群作用とは限らない）
- Contrastive Learning
  元データと変換データは近い出力を，
  他の画像は遠い出力なるよう学習．
- SimCLR [Chen et al 2020],
  Contrastive Predictive Coding [van der Ord et al 2018] etc


(a) Original (b) Crop and resize (c) Crop, resize (and flip) (d) Color distort. (drop) (e) Color distort. (jitter)
(f) Rotate {90°, 180°, 270°} (g) Cutout (h) Gaussian noise (i) Gaussian blur (j) Sobel filtering



群の表現

群の表現　群の線形な作用

- 群の表現

  群の性質を線形変換を通して解析するための数学的方法.

  <u>定義</u>. G: 群. V: ベクトル空間. $GL(V)$: V上の可逆線形変換全体.

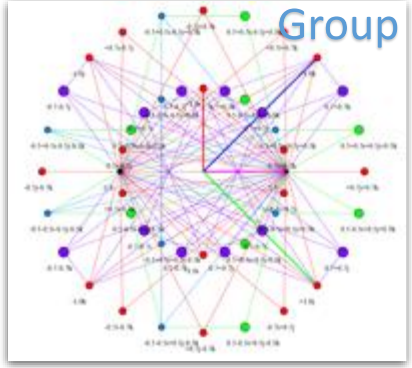
$$\rho : G \to GL(V)$$

が群 G のV上の表現であるとは, ρ が<u>群準同型</u>であること,
すなわち,

$$\rho(e) = I_V, \quad \rho(ab) = \rho(a)\rho(b) \qquad (\forall a, b \in G).$$

ρ はG の V への線形な作用を定める.

$$x \mapsto \rho(a)x$$



$$a \cdot b$$

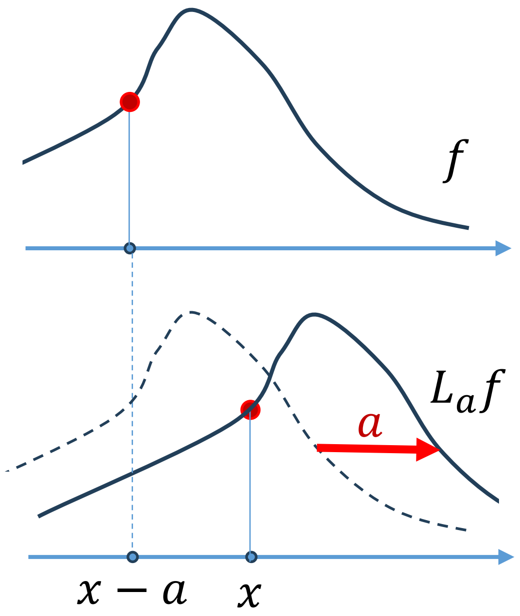
$$\Big\downarrow \rho$$

$$\big(\ \rho(a)\ \big)\big(\ \rho(b)\ \big)$$

行列

正則表現

関数のシフトを表す. CNNで使われている.



定義. G: 群. $V \coloneqq \{f: G \to \mathbb{R}\}$ G上の関数全体.

$$L_g: V \to V,\ \ f \mapsto [h \mapsto f(g^{-1}h)]$$

L_g はGの表現（正則表現と呼ぶ），
すなわち
- 線形写像,　　（$L_g(c_1 f_1 + c_2 f_2) = (c_1 f_1 + c_2 f_2)(g^{-1}\cdot) = c_1 L_g(f_1) + c_2 L_g(f_2)$）
- $L_{gh} = L_g \circ L_h$.

例）　$G = (\mathbb{R}^d, +)$, $a \in G$.　$(L_a f)(x) = f(-a + x) = f(x - a)$.

群表現の分解

定義（既約表現）群G の表現 (ρ, V) に対し，$\{0\}$でもVでもない部分空間 $W \subset V$ があって $\rho(g)W \subset W$を満たす（したがってWへの制約がまた表現になる） とき， (ρ, V) は可約 であるという．可約でないとき既約という．

可約 $$\begin{pmatrix} A(g) & B(g) \\ O & D(g) \end{pmatrix} \begin{pmatrix} w \\ 0 \end{pmatrix} = \begin{pmatrix} A(g)w \\ 0 \end{pmatrix}$$

定義（直和） 群G の表現 $(\rho_1, V_1), (\rho_2, V_2)$ の直和：

$$\rho_1 \oplus \rho_2 : G \to V_1 \oplus V_2,$$
$$g \mapsto \rho_1(g) \oplus \rho_2(g).$$

$$\rho(g) = \begin{pmatrix} \rho_1(g) & O \\ O & \rho_2(g) \end{pmatrix}$$

- 定義　表現 ρ が完全可約（半単純）とは，既約表現の直和と同値になること．

$$\rho \cong \rho_1 \oplus \rho_2 \oplus \cdots \oplus \rho_k$$

共通の基底をとって，任意のgに対して同時ブロック対角化できる場合．

$$\rho(g) = P \begin{pmatrix} A_1(g) & O & \cdots & O \\ O & A_2(g) & \ddots & O \\ \vdots & & & \vdots \\ O & & \cdots & A_k(g) \end{pmatrix} P^{-1}$$

基底変換 P は gに非依存
各$A_i(g)$ は既約表現

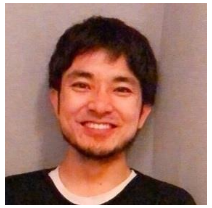
- 以下のクラスは，任意の有限次元表現が完全可約となる．
  - 有限群
  - 局所コンパクトアーベル群 ($SO(2)$, etc)
  - コンパクトLie群 ($O(n)$, etc).
- 完全可約でない例：　ユークリッド運動群 E(n)

$$E(n) = \left\{ \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \middle| A \in O(n), b \in \mathbb{R}^n \right\}$$

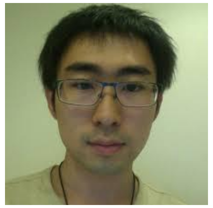
$$\begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Ax + b \\ 1 \end{pmatrix}$$

Neural Fourier Transform:
データからの学習によるFourier変換

Miyato, Koyama, Fukumizu. NeurIPS 2022;
Koyama, Fukumizu., Hayashi, Miyato. ICLR 2024



Takeru Miyato
(U Tubingen/
Preferred Networks)

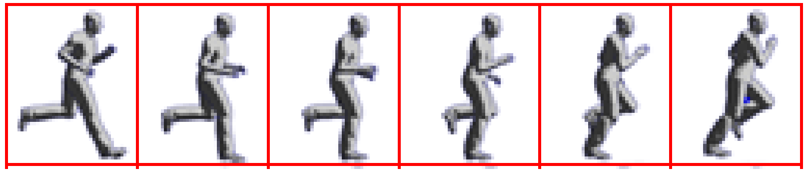


Masanori Koyama
(Preferred Elements)

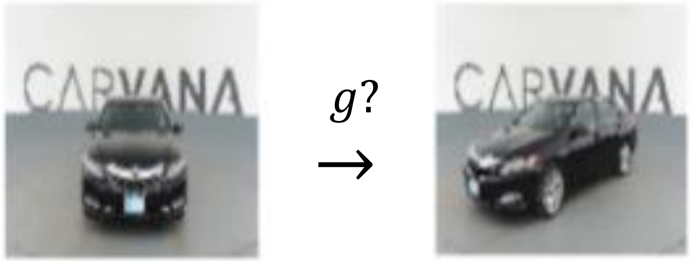
Kohei Hayashi
(Preferred Elements)

群作用を用いた系列データからの表現学習

(Miyato, Koyama, Fukumizu, NeurIPS 2022; Koyama, Fukumizu, Hayashi, Miyato, ICLR 2024)



- 潜在的な群作用を用いた表現学習
  - 作業仮説：「動き」の中に有効な特徴が潜在している
  - **群作用**から生じると仮定する
    ペア $(x, g \circ x)$,
    系列 $(x, g \circ x, g^2 \circ x, \ldots, g^T \circ x)$
    からの表現学習
  - 潜在的な群作用．群は知らない
    - 画像自体への変換とは限らない
    - 直接観測できるとは限らない



3D空間での回転

非線形な観測（魚眼レンズ）

データからの未知/既知の対称性の抽出

最も基本の問題設定：教師なしの表現学習

- データの生成モデル：群作用による定常な系列データ

  G: 群，$\mathcal{S}$ に作用.

  系列： $\mathbf{s} = (s_0, s_1, s_2, \dots, s_T)$.

  $$s_t = g^t \circ s_0, \quad \mathbf{s} = \mathbf{s}(g, s_0)$$

  $g \in G$ は系列ごとに異なってよいが系列内では一定.

- 多数の系列データからの教師無し学習

  G も $g \in G$ も作用も知らない

$\mathbf{s}^{(1)} = \mathbf{s}\left(g^{(1)}, s_0^{(1)}\right)$　　e.g. $g^{(1)} = (15^\circ$回転,（-2,1)シフト，ΔRGB)

$\mathbf{s}^{(2)} = \mathbf{s}\left(g^{(2)}, s_0^{(2)}\right)$

回転，平行移動，色の回転

初期画像 s_0 と群の元 $g \in G$ を決めると，ひとつの系列 $\mathbf{s}$ が生成される.

Autoencoderによる線形な遷移則の学習

Meta Sequential Learning （MSP)

$\mathbf{s} = (s_t)_{t=0}^{T}, \quad s_{t+1} = g \circ s_t$

- Autoencoder：

  Encoder $\Phi: \mathcal{X} \to \mathbb{R}^{a \times m}$, Decoder $\Psi: \mathbb{R}^{a \times m} \to \mathcal{X}$

- 潜在表現の遷移則は**線形**：

  行列 $M_{\mathbf{s}}$ 系列依存.

$$\begin{array}{cc} & m \\ \begin{pmatrix} M_{\mathbf{s}} \end{pmatrix} \begin{pmatrix} \quad \Phi(s) \quad \end{pmatrix} & a \end{array}$$

- 潜在空間は行列：

  表現能力を落とさずに $M_{\mathbf{s}}$ の次元を小さくできる.

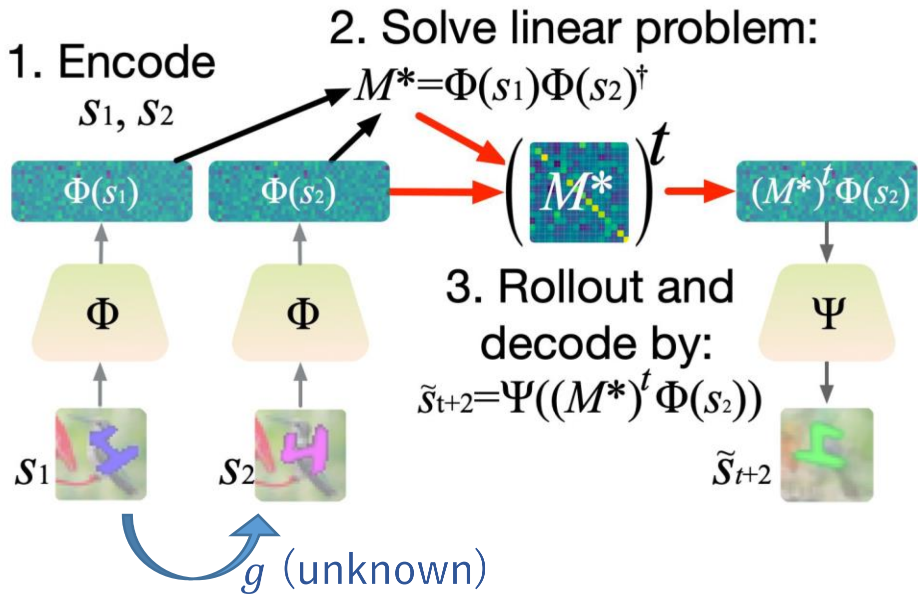
- $M_{\mathbf{s}}, \Phi, \Psi$ の学習

$$\min_{M_{\mathbf{s}}} \| M_{\mathbf{s}} \Phi(s_t) - \Phi(s_{t+1}) \|^2 \qquad \text{[同変性]}$$

$$\min_{\Phi,\Psi} \textstyle\sum_{\ell} \left\| \Psi\left(M_{\mathbf{s}}^{\ell} \Phi(s_t) \right) - s_{t+\ell} \right\|^2 \qquad \text{[Prediction]}$$

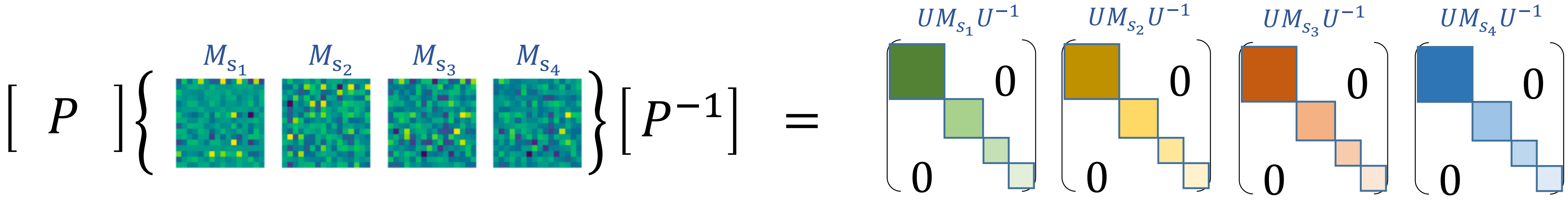
Φを固定すると,
最小２乗誤差でMが決まる.

モデル



変換行列の既約分解

- 変換行列 $M_{\mathbf{s}}$ は系列 $\mathbf{s}$ ごとに異なる
  変換 g に依存，$M(g)$と書ければ群の表現
- 表現の既約分解 → $\{M_{\mathbf{s}}\}_{\mathbf{s}}$ の同時ブロック対角化
  系列に依存しない基底変換により，同時にブロック対角化を行う
  各ブロックは既約表現 → 分解された表現.



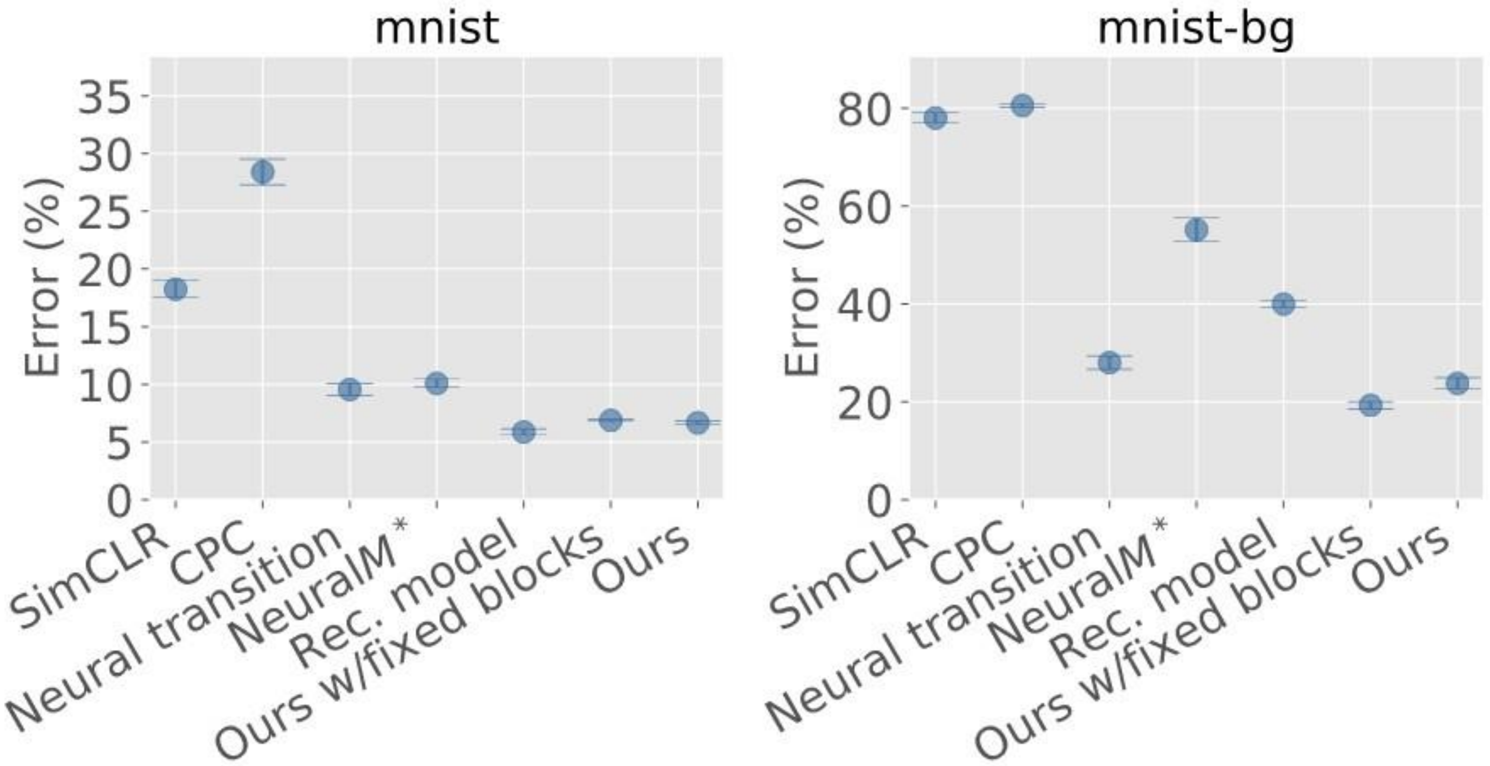
基底変換 P は共通

表現論によるDisentanglement

実験 1: 表現学習としての有効性

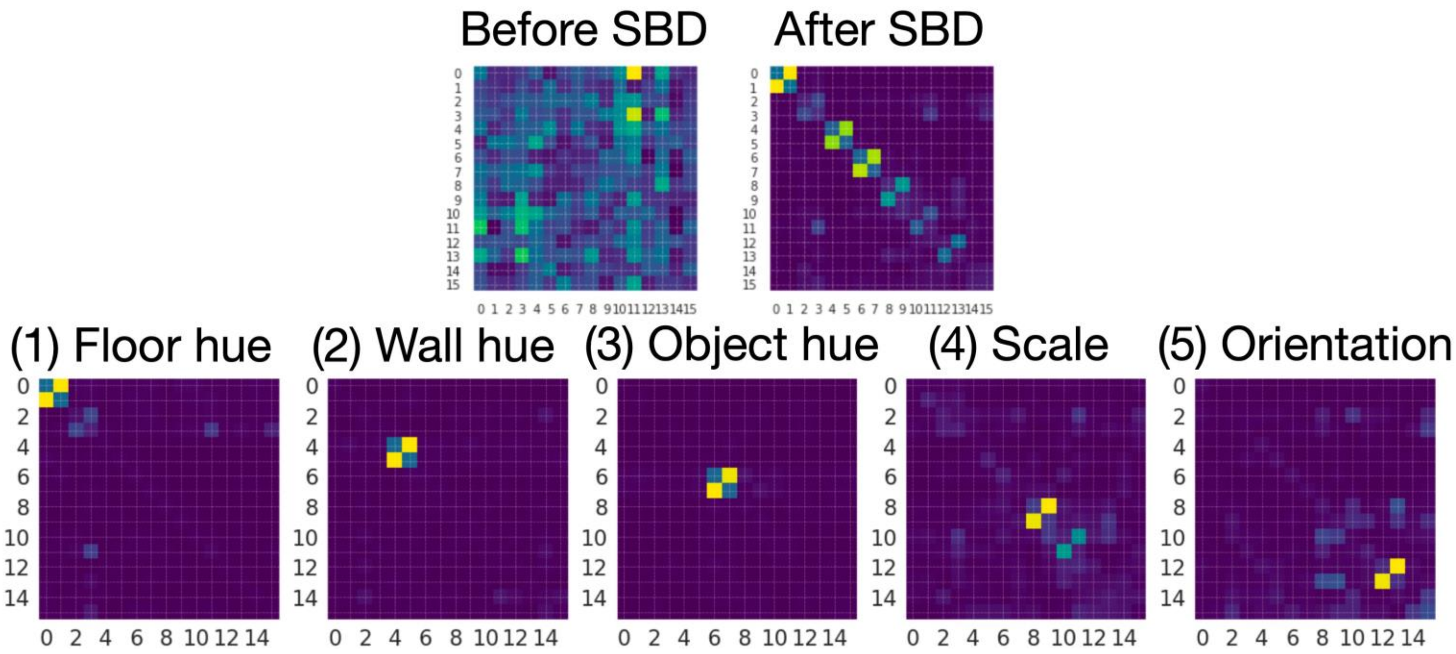
- Encoder出力 $\Phi(x)$ を用いて<u>線形</u>識別タスクを解く

  MNIST 画像“4”の系列データのみで学習 → “0”,⋯,”9”. 10クラスの識別を解く.

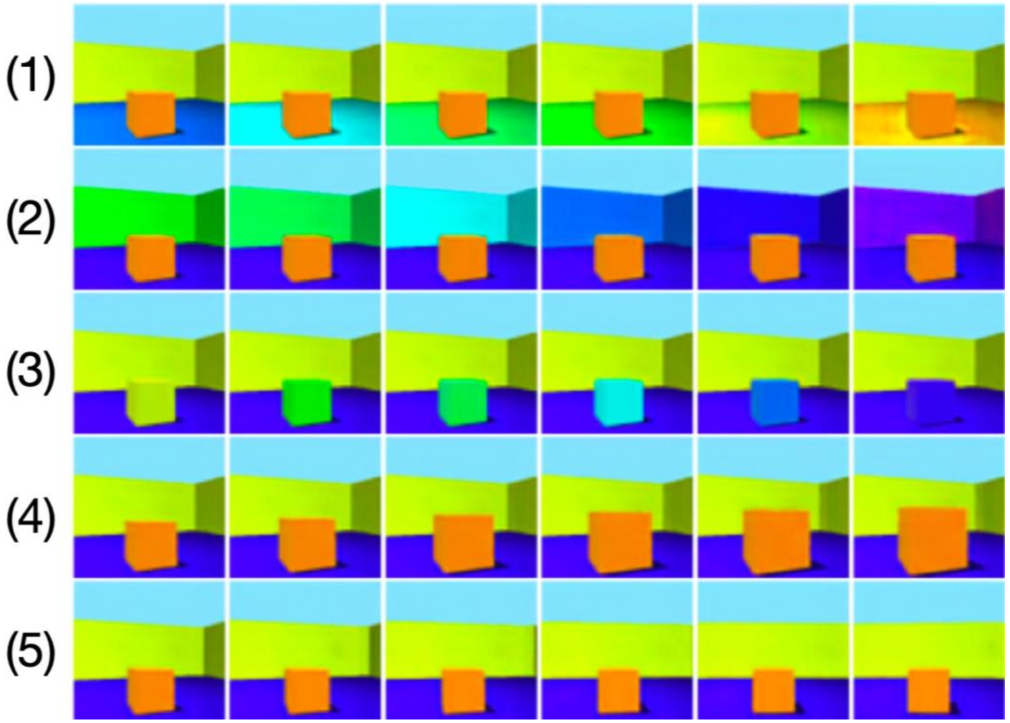


* SimCLR, CPC: 自己教師あり学習

• 実験２：既約分解によるDisentanglement



同時ブロック対角化により得られたブロック



各ブロックのみを取り出した M による生成結果

Fourier変換

- （復習）古典的なFourier変換
  $\mathbb{S}^1 \cong [0,1]$ 上のFourier変換 (Fourier級数展開)
  $$\hat{f}_n = \Phi(f)(n) = \int_0^1 e^{-2\pi\sqrt{-1}nx} f(x)dx, \qquad f(x) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{2\pi\sqrt{-1}nx}\hat{f}_n$$
  $\mathbb{S}^1$は加法群，fは群上の関数

- Fourier変換はより一般の群に対して定義可能
  - 可換群（局所コンパクトアーベル群）：　ほぼそのまま拡張可能
    $$\Phi : L^2(G) \to L^2(\hat{G}), \quad f \mapsto \hat{f}(\rho) := \int_G \overline{\rho(x)} f(x)dx.$$
    $\rho(x)$ は既約表現（常に1次元）（$\mathbb{S}^1$の例では，ある n に対する $e^{2\pi\sqrt{-1}nx}$ に相当）

- コンパクト群
  Fourier変換

$$\Phi(f) = \hat{f}(\rho) := \int_G f(g)\rho(g)^* dg = \int_G f(g)\rho(g^{-1}) dg$$

$[\rho] \in \hat{G} := \{G$の既約ユニタリ表現の同値類$\}$ 周波数の空間

   - $\rho(g)$ は一般に行列値関数. , $d_\rho := \dim \rho$

- Fourier inversion

$$f(g) = \sum_{[\rho]\in\hat{G}} d_\rho \mathrm{Tr}\left[\rho(g)\hat{f}(\rho)\right]$$

Fourier変換の同変性

- Fourier変換の同変性
  - $\mathbb{S}^1 \cong [0,1]$上のシフト $L_a f \coloneqq f(\cdot - a)$ に対するFourier変換

$$\Phi(L_a f) = \left(e^{2\pi\sqrt{-1}an}\hat{f}_n\right) = \begin{pmatrix} \ddots & \cdots & 0 \\ \vdots & e^{2\pi\sqrt{-1}a\,n} & \vdots \\ 0 & \cdots & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ \hat{f}_n \\ \vdots \end{pmatrix}$$

$$= M(a)\Phi(f)$$

$$\begin{tikzcd} f \arrow[r, "L_a"] \arrow[d, "\Phi"'] \arrow[dr, phantom, "\circlearrowright"] & f(\cdot - a) \arrow[d, "\Phi"] \\ (\hat{f}_n) \arrow[r, "M(a)"'] & M(a)(\hat{f}_n) \end{tikzcd}$$

- $\mathbb{S}^1 \ni a \mapsto M(a)$ は加法群 $\mathbb{S}^1$の表現（線形作用）
- Φ は同変（右の図式が可換）
- 同変性は一般のFourier変換に対して成り立つ.

Neural Fourier Transform

- 提案法はFourier変換の拡張と考えられる

群 Gが空間 $\mathcal{X}$ に作用しているとき，以下の写像をデータから学習している．

ブロック対角（既約表現）

$$L_g x \approx \Psi \circ M(g) \circ \Phi(x) = \underbrace{\Psi \circ P^{-1}}_{\text{Fourier逆変換}} \circ \begin{pmatrix} \rho_1(g) & & O \\ & \rho_2(g) & \\ O & & \ddots \end{pmatrix} \circ \underbrace{P \circ \Phi(x)}_{\text{Fourier変換}}.$$

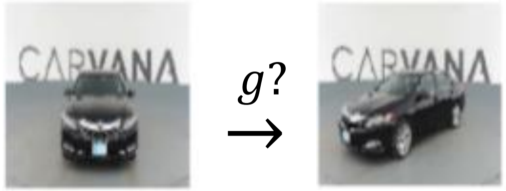
$$x \approx \Psi \circ P^{-1} \circ P \circ \Phi(x)$$ Fourier inversion formula

Neural Fourier Transform:
データからの学習による，任意のデータ型に対するFourier変換

- 「潜在的空間」への群の作用でOK.
- 群上の関数ではなく，データに対するFourier変換
- データの記述に必要な周波数/既約成分のみを自動獲得

- さまざまなシナリオ

	群 G	群の元 g	作用	$M(g)$の推定
U-NFT	Unknown	Unknown	Unknown	MSPで学習
G-NFT	Known	Unknown	Unknown	調和関数の関数系利用
g-NFT	Known	Known	unknown	調和関数 Φ, Ψのみ学習

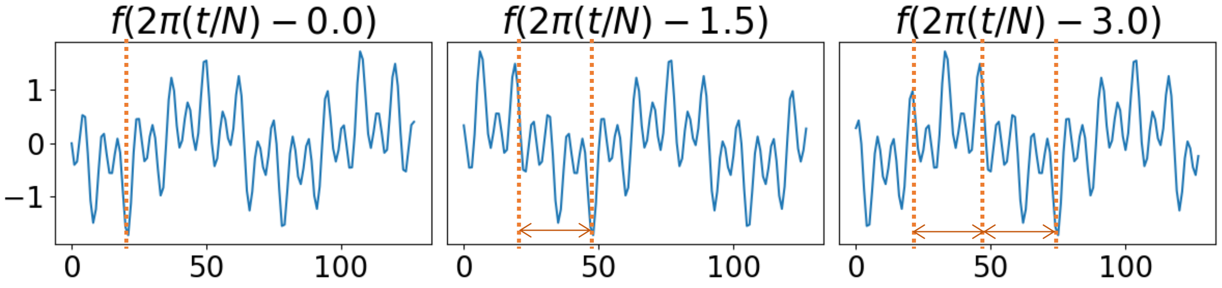


• 未知の非線形変換による観測（魚眼レンズ）

潜在空間（一定のシフト）観測されない

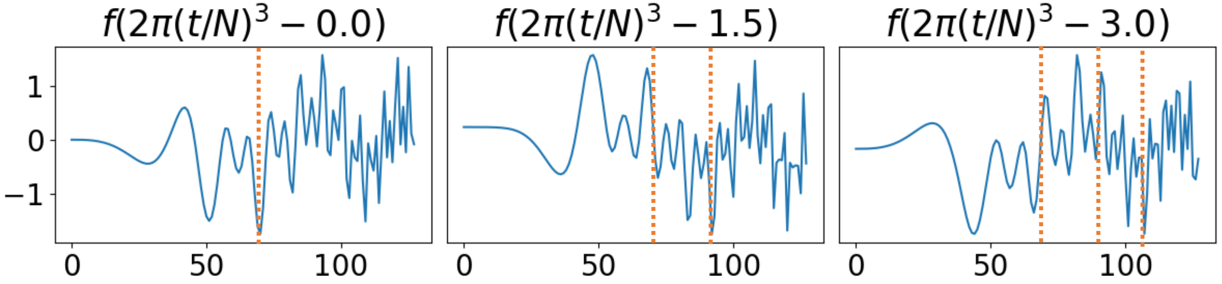
- １次元信号の場合

周波数５個の混合信号

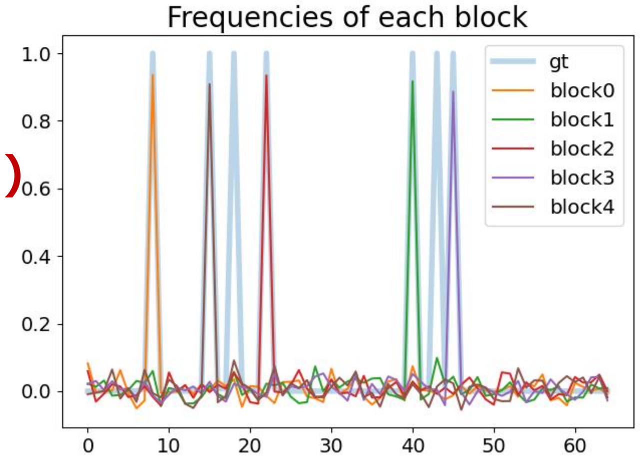


未知の非線形観測 $\left(\frac{t}{128}\right)^3$

非線形観測信号



U-NFT（MSPで学習）



正しい周波数を推定

- ２次元画像の場合

予測画像系列

真の系列

- g-NFT：2D画像1枚からの3次元生成

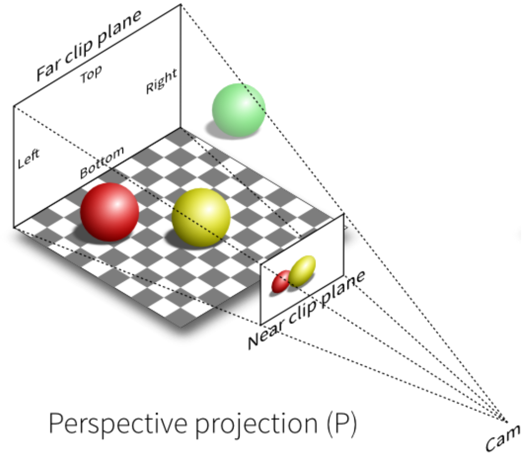
学習： 3D-CGデータを観測した２D画像データのみによるg-NFT学習

2D画像ペアと３D回転データ (S, S', R) $R \in SO(3)$, S, S': 2D画像

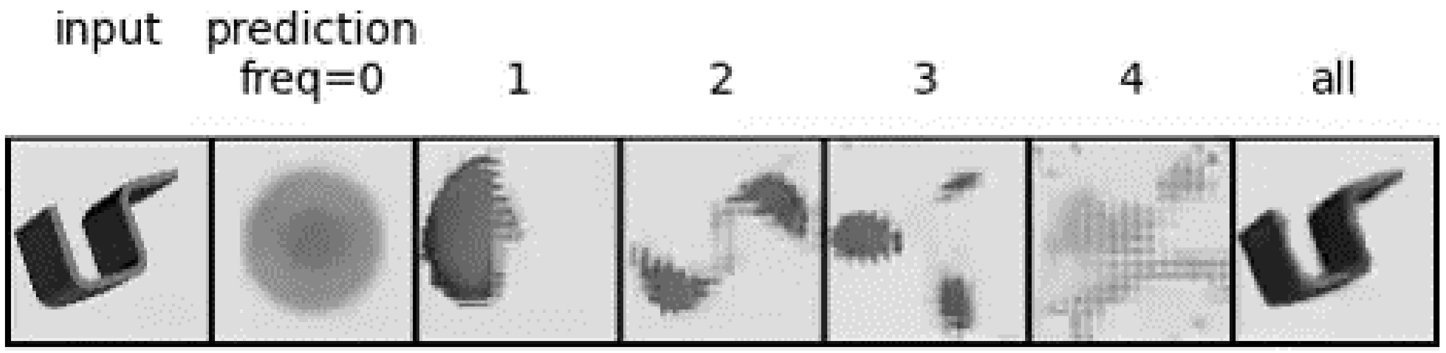
$M(g)$ 球面調和関数を仮定．Encoder, decoderのみ学習

使用時： 訓練データにない２D画像１枚

→ 任意の３次元回転を施した画像を生成



Perspective projection (P)



まとめ

- 群作用を利用する深層学習
  - アーキテクチャ: G-CNN
  - データ拡張：　不変学習，自己教師あり学習
  - 同変性による表現学習：　新しい方向性
- Neural Fourier Transform
  - 潜在空間で線形作用になるような学習
  - 群の作用を知らなくても適用可能
  - データ駆動の非線形Fourier変換が実現可能
    - 任意のデータ型に対するFourier変換
    - 必要な周波数を自動抽出
  - 応用（今後）
    - シミュレーションで学習 → 静止２D画像の物体を自由に動かす
    - 変微分法方程式系の特徴学習

References

- Online materials
  - U Amasterdam - An Introduction to Group Equivariant Deep Learning https://uvagedl.github.io/

- Textbooks on Group Theory
  - [Group Theory] J.J. Rothman. An Introduction to the Theory of Groups. Springer 1995.
  - [Representation Theory] W. Fulton, J. Harris. Representation Theory. Springer 2004
  - [Fourier Analysis] W. Rudin. Fourier Analysis on Groups. John Wiley& Sons, 1991

- Papers

  - Fukushima, Kunihiko (1980). "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position". *Biological Cybernetics* **36** (4): 193–202. doi:10.1007/BF00344251
  - Zhang, Wei (1988). "Shift-invariant pattern recognition neural network and its optical architecture". *Proceedings of annual conference of the Japan Society of Applied Physics*.
  - Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, Backpropagation Applied to Handwritten Zip Code Recognition Archived 2020-01-10 at the Wayback Machine; AT&T Bell Laboratories
  - T. Cohen and M. Welling. Group Equivariant Convolutional Networks. ICML 2016
  - Lafarge, M. W., Bekkers, E. J., Pluim, J. P. W., Duits, R., & Veta, M. (2021). Roto-translation equivariant convolutional networks: Application to histopathology image analysis. *Medical Image Analysis*, *68*, 101849. https://doi.org/https://doi.org/10.1016/j.media.2020.101849
  - M. Weiler and G. Cesa. General E(2) - Equivariant Steerable CNNs. NeurIPS 2019

- T. Miyato, M. Koyama, K. Fukumizu. Unsupervised Learning of Equivariant Structure from Sequences. NeurIPS 2022
- M. Koyama, K. Fukumizu, K. Hayashi, T. Miyato. Neural Fourier Transform: A General Approach to Equivariant Representation Learning. ICLR 2024 https://openreview.net/forum?id=eOCvA8iwXH
- Thomas W. Mitchel, Michael Taylor, Vincent Sitzmann. Neural Isometries: Taming Transformations for Equivariant ML. arXiv:2405.19296 [cs.CV]
- Jianke Yang, Nima Dehmamy, Robin Walters, Rose Yu. Latent Space Symmetry Discovery. ICML 2024 https://openreview.net/forum?id=xbXASfz8MD

Appendix

- 例: $\mathfrak{S}_3$ の既約表現

  1) 自明な表現 (1 dim)

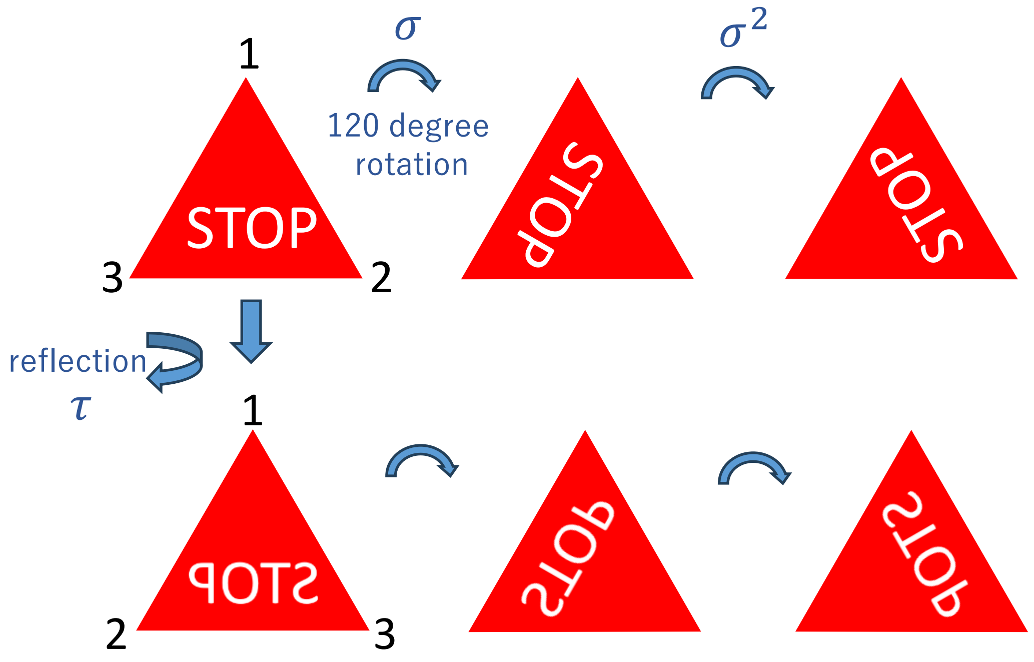
  $$\rho_e(g) = 1$$

  2) Alternating representation (1 dim)

  $$\rho_{alt}(g) = \mathrm{sgn}(g)$$

  3) 標準表現 (2 dim)

  $$\begin{cases} \rho_{st}(\sigma) = \begin{pmatrix} \cos(-2/3\pi) & -\sin(-2/3\pi) \\ \sin(-2/3\pi) & \cos(-2/3\pi) \end{pmatrix} = \begin{pmatrix} -\sqrt{3}/2 & 1/2 \\ -1/2 & -\sqrt{3}/2 \end{pmatrix} \\ \rho_{st}(\tau) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \end{cases}$$



$\mathfrak{S}_3$ の既約表現は上の3つだけであることが知られている.(See, e.g., Fulton Harris)

G-CNNの目的

$$E(n) = \left\{ \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \middle| A \in O(n), b \in \mathbb{R}^n \right\}$$

$$\begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Ax + b \\ 1 \end{pmatrix}$$

- Symmetries
  - 画像の多くの特徴はEuclid運動群（回転＋シフト）に対して不変(E(2), SE(2)).
    - 医療画像データ: 回転やシフトはあまり意味を変えない.

      c.f., 自然画像(部屋, 道路など), 文字画像 (アルファベット, 数字など)は回転で意味が変わる.

  - 3D データ/360度画像（SO(3), E(3)）
    - 3D スキャナ, レンダリング.
    - 360度カメラ.

      https://github.com/QUVA-Lab/e2cnn?tab=readme-ov-file

  - グラフ
    - 置換に関して不変
    - ３次元内のグラフ構造, e.g., 分子（SE(3)）