

# 人工ニューラルネットワークによる 機械学習 ～その発明と応用～

理化学研究所 革新知能統合研究センター（理研AIP）

上田 修功

# 自己紹介

本務) 理化学研究所 革新知能統合研究センター 副センター長  
目的指向基盤技術研究グループディレクター  
iPS細胞連携医療リスク回避チームリーダー  
防災科学チームリーダー

兼務) 日本電信電話 (株) NTTコミュニケーション科学基礎研究所  
リサーチプロフェッサー (NTT客員フェロー)

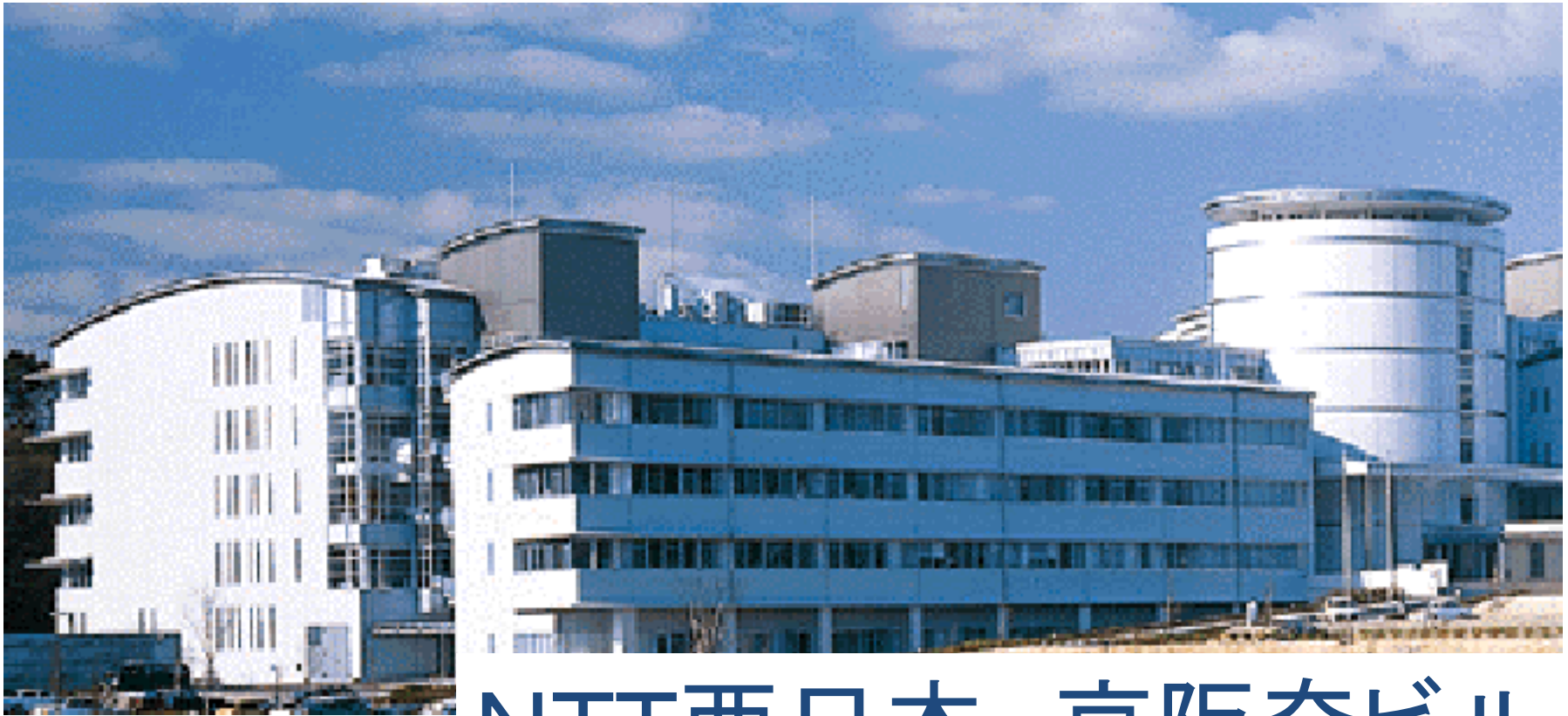
- 日本科学技術振興機構(JST) 数理的情報活用基盤 CREST研究総括
- ムーンショット 目標3 (AI x ロボティクス) AI担当サブPD

研究分野: 統計的機械学習とその自然科学  
(宇宙物理、医療、創薬、地震、気象) への応用



# コミュニケーション科学基礎研究所

1991.7.4誕生



NTT西日本 京阪奈ビル

けいはんな学研都市

(京都府相楽郡精華町光台2-4)



# 人工知能（AI）ブームの変遷

第1次AIブーム（1950～1960年代）

探索 & 推論

第2次AIブーム（1980年代）

知識表現，学習

第3次AIブーム（2000年代）

ビッグデータ＋深層学習

# 演繹法(科学) vs. 帰納法(機械学習)

## 演繹法

前提

一般的、普遍的な事実

正当化

結論

### 演繹的論理展開の例

「ビッグデータを制するものがビジネスを制する」

「機械学習技術はビッグデータ解析に有利」

「企業は機械学習技術が必要としている」

## 帰納法

さまざまな事例

結論

### 帰納的論理展開の例

「A社は機械学習法aでビッグデータ解析をしている」

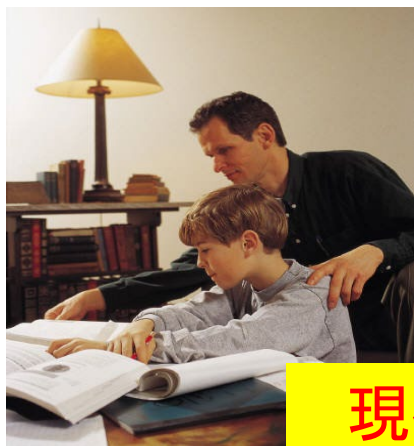
「B社は機械学習法bでビッグデータ解析をしている」

「C社は機械学習法cでビッグデータ解析をしている」

「ビッグデータ解析のために機械学習技術が必要である」

# 学習のタイプ

## 人の学習と似ている



教師なし学習  
(自習する)

ex) 特徴選択、特徴抽出、  
次元圧縮、クラスタリング

現在、深層学習はこれら全てに貢献

教師あり学習  
(先生に習う)

ex) 画像認識, 音声認識,  
言語翻訳などのパターン認識全般

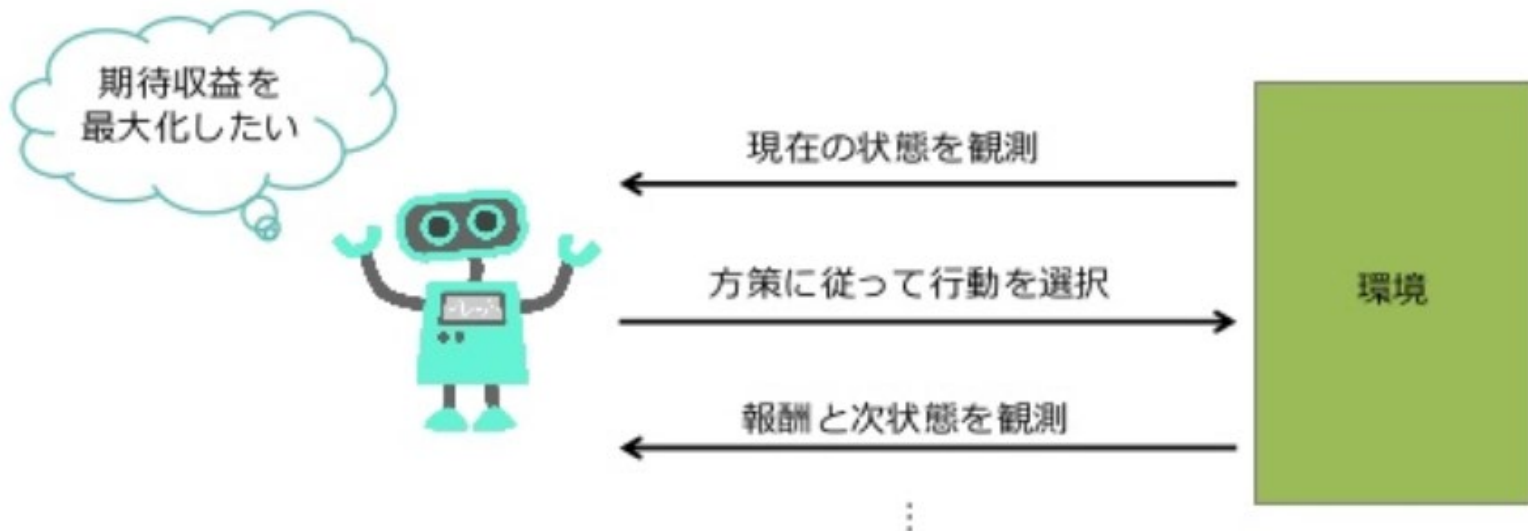


強化学習

(試行錯誤して報酬を最大化する)

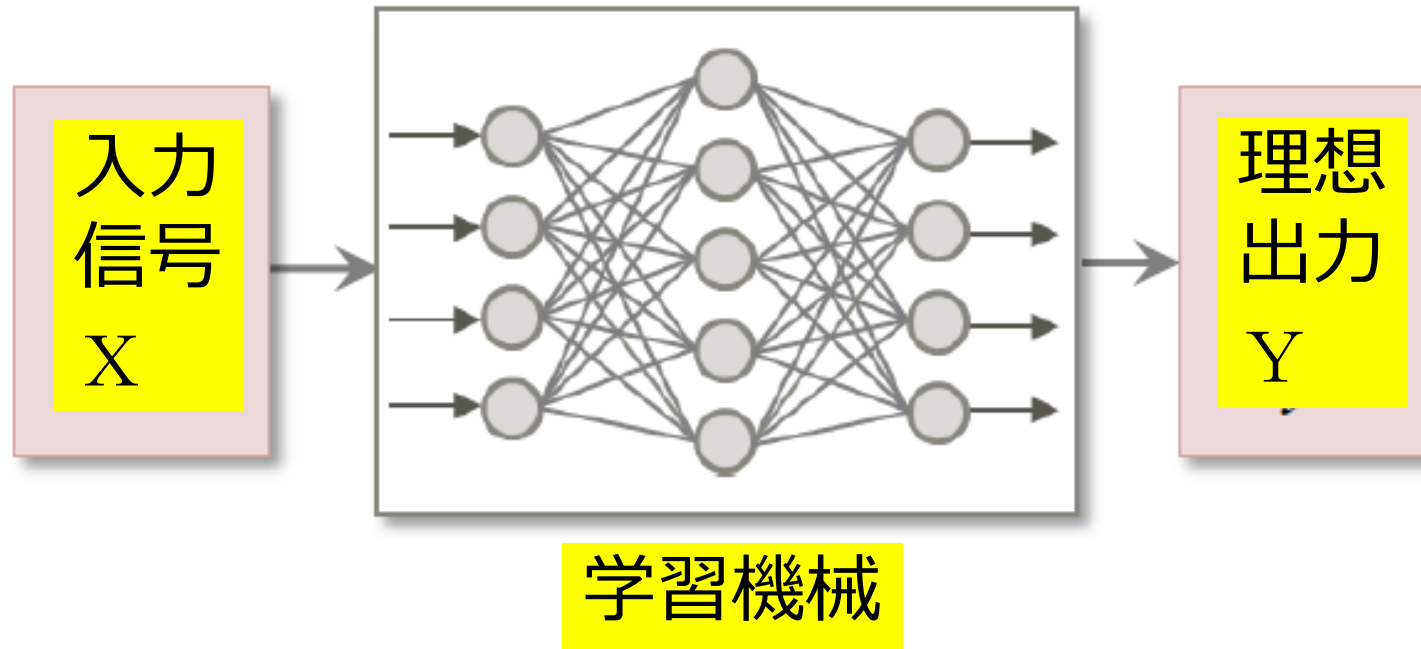
ex) 歩行ロボット, ゲーム(アルファ碁)

# 強化学習



# 教師あり学習（主流）

$$Y=f(X;\Theta)$$



- ・学習データ：  $D = \{x, y\}$   $y$ : 教師データ
- ・学習：  $f(x)$ と理想出力との誤差を最小化する $\Theta$ を求める
- ・テスト： 未知入力 $x'$ に対する学習済の学習機械の出力  
 $y'=f(x';\Theta')$

# 多層ニューラルネットワーク

教師あり学習の汎用学習機械

教師信号

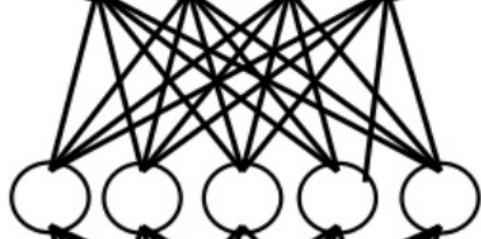


パラメータ更新

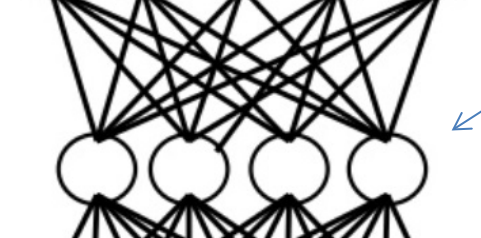
出力層



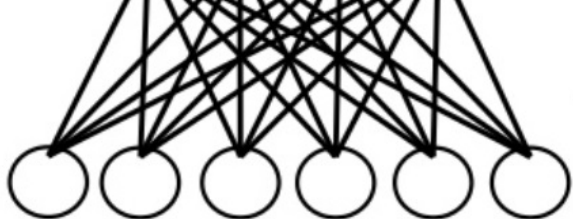
隠れ層



隠れ層



入力層



誤差逆伝搬法

(Back Propagation:BP)

出力と教師信号との誤差を  
最小化するための最急降下法

シグモイド関数



$$f(x_i) = \frac{1}{1 + e^{-x_j}}$$

Rectified Linear Unit(ReLU)



$$f(x_i) = \max(x_j, 0)$$

Maxout



複数のカーネルの畳み込み  
値の最大を出力

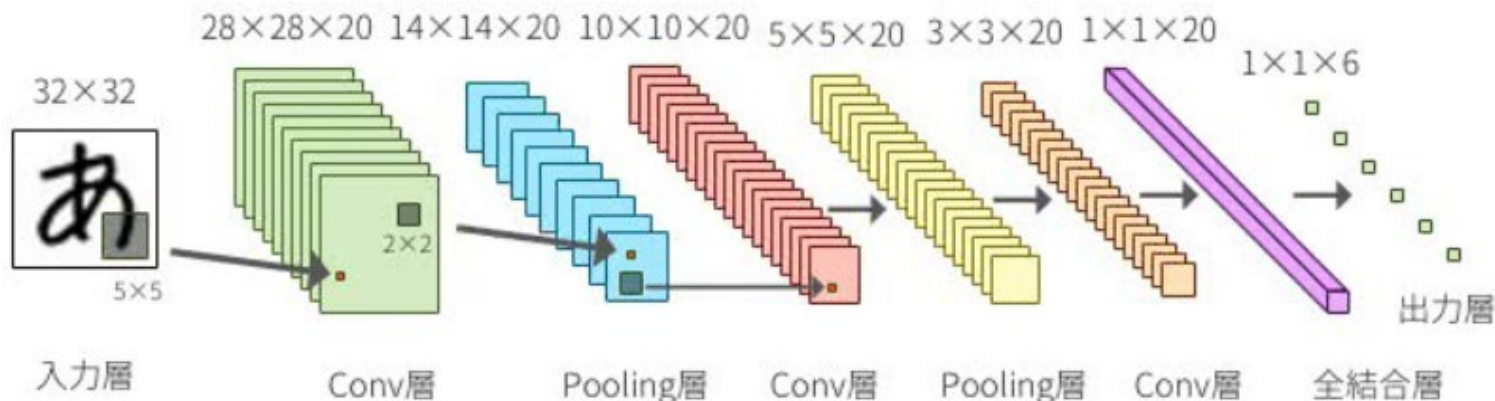
活性化関数

積和&活性化計算の反復で、非線形な特徴抽出  
を繰り返している

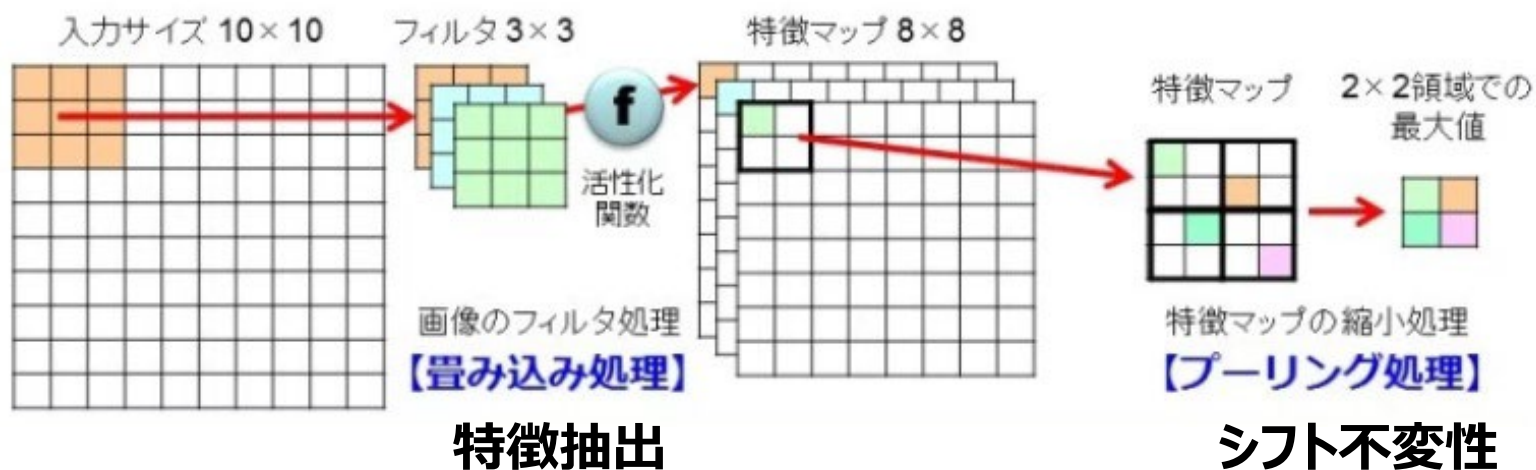


# 畳み込みニューラルネットワーク CNN(Convolutional NN)

人間の視覚系を模擬



畳み込み処理とプーリング処理の反復で異なる解像度での特徴抽出が可能



畳み込みとプーリングの繰り返しにより多様な変形に対処



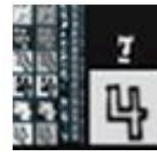
# ニューラルネットワークの変遷



1958 Perceptron  
**F. Rosenblatt**

1986: Nature論文:  
逆誤差伝搬法  
**Rumelhart, Hinton,  
Williams**

1974 Backpropagation  
**P. Werbos**



**CNN Y. LeCun**



Convolution Neural Networks for  
Handwritten Recognition

1998



Google Brain Project on  
16k Cores

2012



**第1次ニューロブーム**

awkward silence (AI Winter)

**第2次ニューロブーム**

**第3次ニューロブーム**

**L.Faifei**

1969

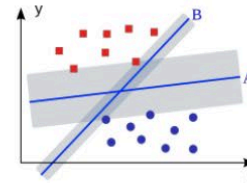
Perceptron criticized



線形分離不能なケースでは限界  
(by M. Minsky: AIの父が  
第1次ニューロブームを終焉させた)

1995

SVM reigns



**V. Vapnik**

2006

Restricted  
Boltzmann  
Machine



**G. Hinton**

2009年、DNNが音声認識において従来の  
state-of-the-art手法の性能を超えた！

2012

AlexNet wins  
ImageNet

IMAGENET

1400万画像、2万  
クラスの良質な  
画像データベース



## 2015年2月には人間の精度を超えた

画像認識で人間の精度を超えることは数十年間、実現されていなかった<sup>16</sup>

Before  
ディープ  
ラーニング

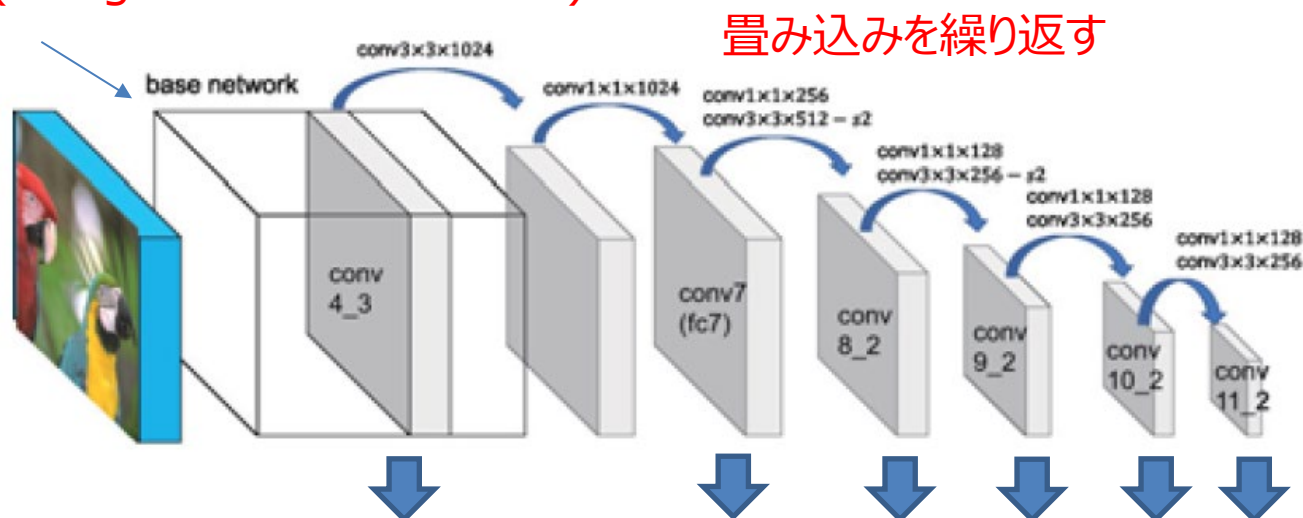
	Error
Imagenet 2011 winner (not CNN)	25.7%
Imagenet 2012 winner	16.4% (Krizhevsky et al.)
Imagenet 2013 winner	11.7% (Zeiler/Clarifai)
Imagenet 2014 winner	6.7% (GoogLeNet)
After ディープ ラーニング	
Baidu Arxiv paper:2015/1/3	6.0%
<b>Human:</b> Andrej Karpathy	5.1%
Microsoft Research Arxiv paper: 2015/2/6	4.9%
Google Arxiv paper: 2015/3/2	4.8%
Microsoft Research CVPR paper: 2015/12/10	3.6%

# SSD: Single Shot Multi-Box Detector

Liu et al., 2015

画像内の物体の検出とクラス分類を同時に高速に行える最新技術

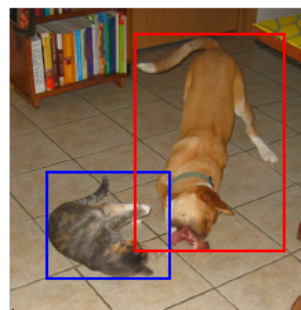
VGG16(ImageNetで学習済のCNN)



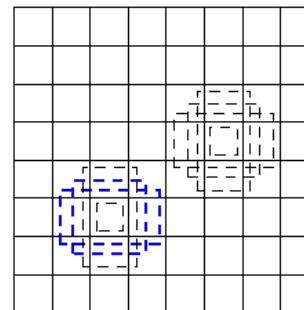
畳み込みを繰り返す

これらの大きさの異なる特徴マップ上で物体の探索とクラス分類を行う

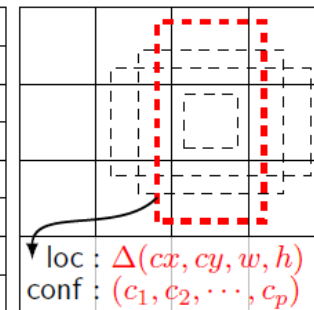
従来、原画像全体を移動窓の  
アスペクト比とサイズを変えながら  
物体探索を行っていたため、計算  
量が膨大となっていたのに対し、  
SSDでは複数の特徴マップによるマ  
ルチスケールの物体探索が可能



(a) Image with GT boxes



(b) 8 × 8 feature map

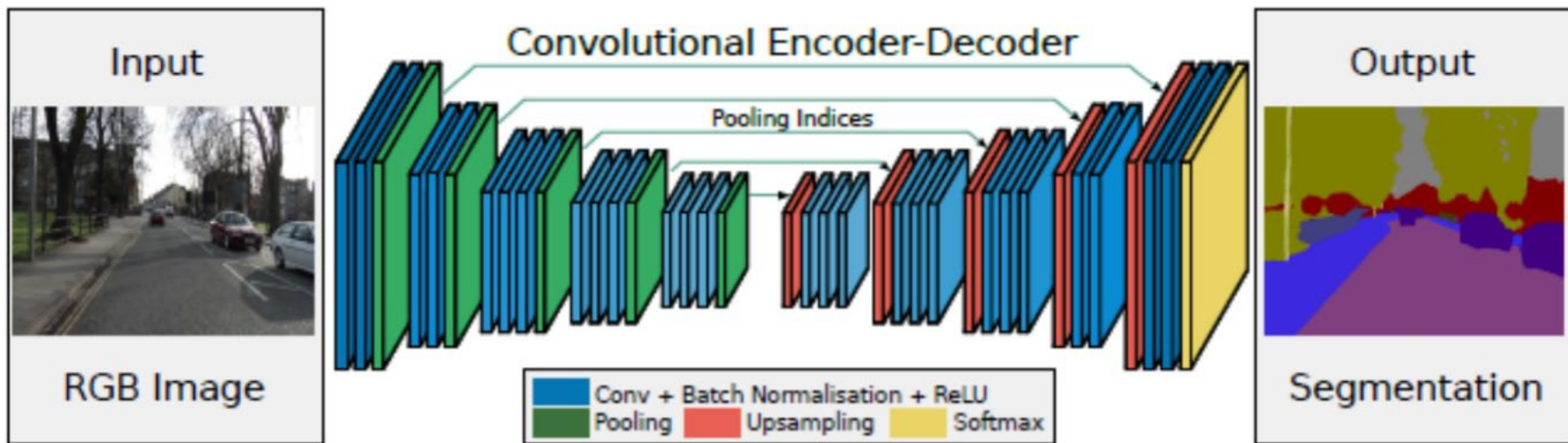


(c) 4 × 4 feature map

Source: <https://arxiv.org/abs/1512.02325> 14

# セマンティックセグメンテーション

SegNet ; 2015



Badrinarayanan et al., arXiv. Preprint arXiv: 1511.0051, 2015.

画像データ → CNN → 画素毎の予測 ↔ 画素毎の正解データ

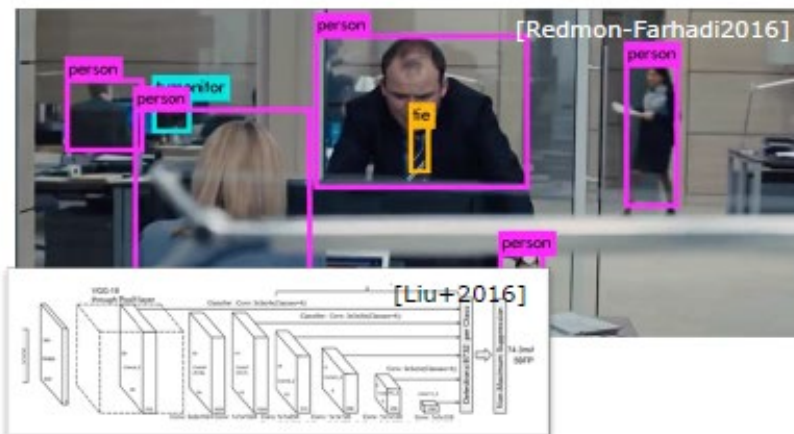
この誤差を最小化するようにCNNを学習

近年、UNET, PSPNetなどの改良手法も提案されている



# 認識

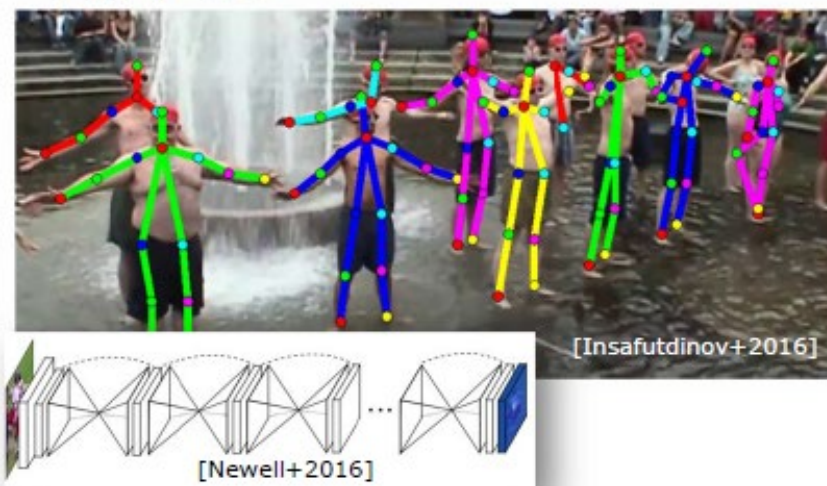
## 物体検出



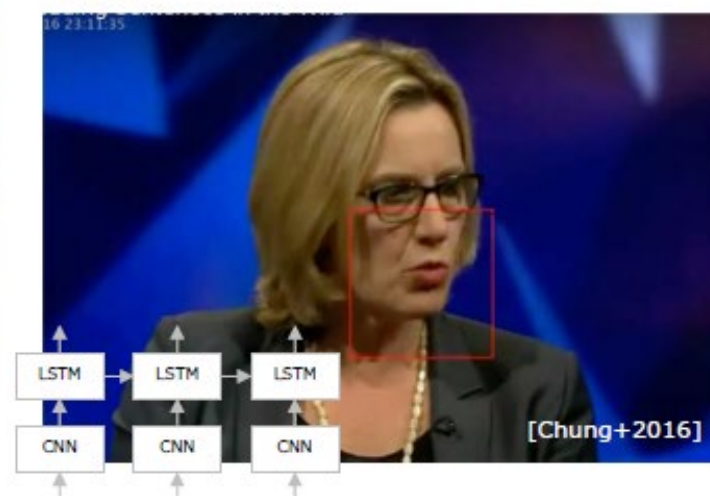
## 画素レベルの認識



## 人体ポーズ



## 読唇



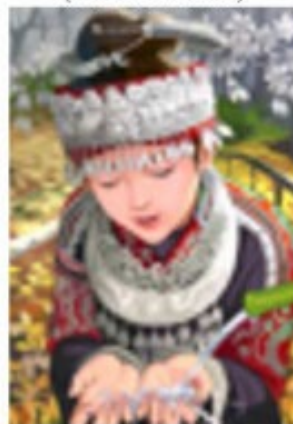
# 画像変換・合成

## スタイル変換

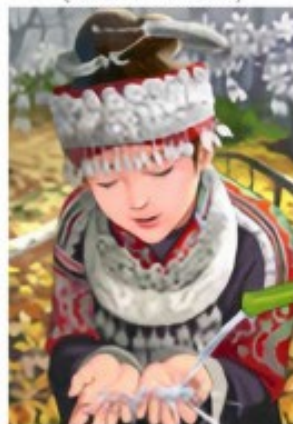


## 超解像

bicubic  
(21.59dB/0.6423)



SRResNet  
(23.53dB/0.7832)



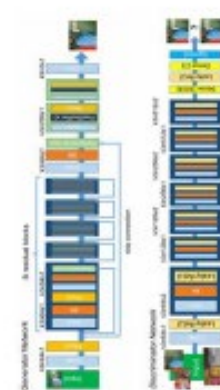
SRGAN  
(21.15dB/0.6868)



original



## 色付け・昼夜変換・線画→実写





# 画像修復 (image inpainting)

Iizuka+, Globally and Locally Consistent Image Completion, SIGGRAPH2017





# Visual Question Answering



**Q: What is reflected in the mirror?**

**A: Cat**



**Q: What room is this?**

**A: Bathroom**



**Q: What is the bird standing on?**

**A: Branch**



**Q: Are these flowers real?**

**A: Yes**



Yoshua Bengio



Geoffrey E Hinton



Yann LeCun



“Artificial intelligence is now one of the fastest-growing areas in all of science and one of the most talked-about topics in society,” said ACM President Cherri M. Pancake. “The growth of and interest in AI is due, in no small part, to the recent advances in deep learning for which Bengio, Hinton and LeCun laid the foundation. These technologies are used by billions of people. Anyone who has a smartphone in their pocket can tangibly experience advances in natural language processing and computer vision that were not possible just 10 years ago. In addition to the products we use every day, new advances in deep learning have given scientists powerful new tools—in areas ranging from medicine, to astronomy, to materials science.”

“Deep neural networks are responsible for some of the greatest advances in modern computer science, helping make substantial progress on long-standing problems in computer vision, speech recognition, and natural language understanding,” said Jeff Dean, Google Senior Fellow and SVP, Google AI. “At the heart of this progress are fundamental techniques developed starting more than 30 years ago by this year's Turing Award winners, Yoshua Bengio, Geoffrey Hinton, and Yann LeCun. By dramatically improving the ability of computers to make sense of the world, deep neural networks are changing not just the field of computing, but nearly every field of science and human endeavor.”

# 生成AIの誕生!

テキスト生成 ChatGPT

画像生成 DALL-E3, Stable Diffusion

動画生成 DALL-E3, Stable Diffusion

音声生成 Voicebox



# 日本の医師国家試験でGPT-4が合格！

## Evaluating GPT-4 and ChatGPT on Japanese Medical Licensing Examinations

Jungo Kasai<sup>♣</sup> Yuhei Kasai<sup>♡</sup> Keisuke Sakaguchi<sup>♣</sup>  
Yutaro Yamada<sup>◇</sup> Dragomir Radev<sup>◇</sup>

<sup>♣</sup>Paul G. Allen School of Computer Science & Engineering, University of Washington

<sup>♡</sup>Sapporo Cardiovascular Clinic <sup>♣</sup>Tohoku University <sup>◇</sup>Yale University

jkasai@cs.washington.edu

必須択一 一般択一 禁忌肢

Model	2018			2019			2020			2021			2022			2023		
	Req.	Gen.	P.↓	Req.	Gen.	P.↓	Req.	Gen.	P.↓	Req.	Gen.	P.↓	Req.	Gen.	P.↓	Req.	Gen.	P.↓
ChatGPT	123	143	1	100	150	5	118	148	2	143	154	3	124	163	2	120	140	–
ChatGPT-EN	123	158	2	117	157	3	116	147	2	110	167	0	140	187	1	142	159	–
GPT-3	105	104	5	93	117	5	97	111	4	94	109	3	106	111	6	86	113	–
GPT-4	161	221	0	170	215	1	168	219	0	173	225	0	164	228	1	170	221	–
Student Majority	196	276	0	196	274	0	195	276	0	200	277	0	195	287	0	–	–	–
Total	200	299	33	200	296	40	197	299	26	200	300	26	197	297	26	200	295	–
Passing Score	160	208	3	160	209	3	158	217	3	160	209	3	157	214	3	160	220	–

arXiv:2023.18027v2[cs.CL] 5 Apr 2023

40歳の初妊婦（1妊0産）。尿糖が陽性であったため、自宅近くの産科診療所から紹介され受診した。現在、妊娠30週。家族歴、既往歴に特記すべきことはない。身長160cm、体重62kg（妊娠前体重55kg）。体温36.7℃。脈拍88/分、整。血圧110/80mmHg。経口グルコース負荷試験〈75gOGTT〉：負荷前値：90mg/dL、1時間値：190mg/dL、2時間値：160mg/dL。HbA1c 5.4%（基準4.6～6.2）。適切な対応はどれか。

*A 40-year-old primigravida (pregnant for the first time, with no prior births) was referred to a nearby obstetrics clinic due to a positive urine glucose test. She is currently at 30 weeks of gestation. There is no significant family or medical history. Her height is 160cm, her weight is 62kg (pre-pregnancy weight was 55kg), and her body temperature is 36.7°C. Her pulse rate is 88 bpm and regular, and her blood pressure is 110/80mmHg. Her 75g oral glucose tolerance test (OGTT) results were as follows: fasting glucose level of 90mg/dL, 1-hour value of 190mg/dL, and 2-hour value of 160mg/dL. Her HbA1c level is 5.4% (normal range 4.6-6.2%). What is the appropriate response?*

a 対応は不要である。 *No response is necessary.*

b 食事は4～6分割食を推奨する。 *Recommend dividing meals into 4-6 portions.* **Correct Answer**

c 食後2時間の血糖値150mg/dLを目標とする。 *Aim for a postprandial blood sugar level of 150mg/dL.*

d 1日の摂取エネルギーを1,200kcalに制限する。 *Restrict daily energy intake to 1,200kcal.*

e 食事療法が無効な場合は経口血糖降下薬を用いる。 *Use oral hypoglycemic agents if dietary therapy is ineffective.*

**ChatGPT's Choice "禁忌 (Prohibited)"**

LLMが医学、金融、法学など、専門的で利害の大きい用途に適用される場合の重大な課題を示している

# 画像理解 GPT4V (Open AI)

Yang(Microsoft)+, The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision), arXiv2023.9

## Sec. 4.4 Scene Text, Table, Chart, and Document Reasoning

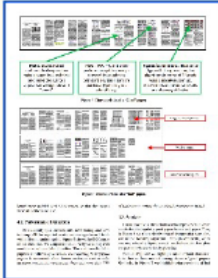
### Document Understanding

#### Prompt:

This is a paper "Paper Gestalt."  
Describe the paper in details  
and highlight their contribution.

Prompt Continued

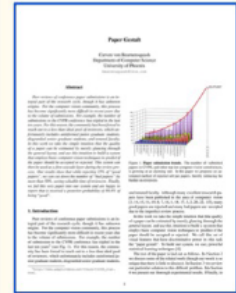
Page 4



Page 5



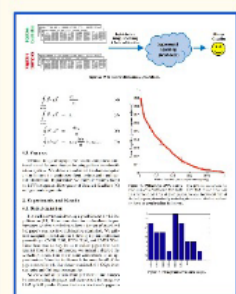
Page 1



Page 2



Page 3



#### GPT-4V:

The paper "Paper Gestalt" is a research paper written by Carven von Beamersquash from the Department of Computer Science at the University of Phoenix. The paper discusses the visual quality of a computer science research paper and how it affects the overall impression of the paper.

The paper begins with an introduction to the concept of paper gestalt, which is the overall impression of a paper. The author argues that the visual quality of a paper is an important factor in creating a good impression. The paper then discusses previous work in this area and highlights the contribution of this paper, which is the use of a supervised learning algorithm to evaluate the visual quality of a paper.

The paper then presents the methodology and experimental results. The author collected a dataset of 1196 computer science research papers and used a supervised learning algorithm to evaluate the visual quality of the papers. The algorithm was trained on a set of features extracted from the papers, such as the number of figures, the number of equations, and the number of references.

algorithm was  
visual quality  
of accuracy.

discussion of  
the author

suggests that future work could focus on improving the algorithm by incorporating more features and training it on a larger dataset. The paper also suggests that the algorithm could be used to evaluate the visual quality of other types of documents, such as technical reports and dissertations.

Overall, the contribution of this paper is the use of a supervised learning algorithm to evaluate the visual quality of a computer science research paper. This is a novel approach that could help researchers improve the visual quality of their papers and make a better impression on their readers.

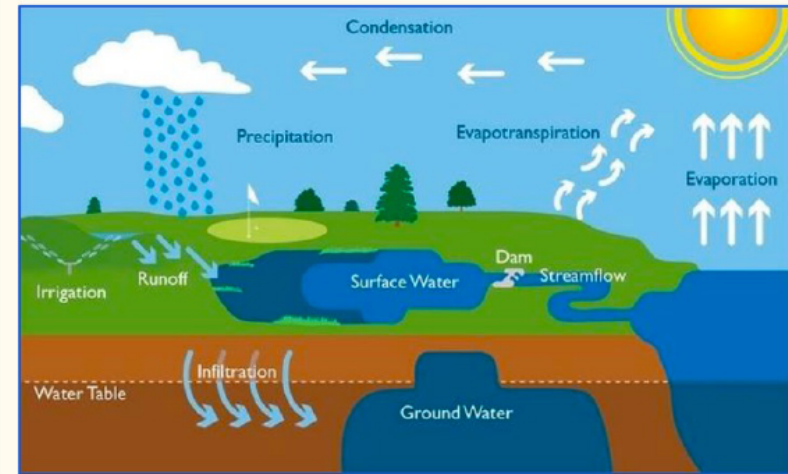
論文の画像 → 論文を要約

## Sec. 4.3 Multimodal Knowledge and Commonsense

### Science and Knowledge

#### Prompt:

Suppose you are a teacher, please use the figure to explain the distinction between evaporation and evapotranspiration.



#### GPT-4V:

Okay class, today we are going to learn about the distinction between evaporation and evapotranspiration. Let's take a look at this figure.

First, let's understand what evaporation is. Evaporation is the process by which water is changed from its liquid state to its gaseous state. This usually takes place at the surface of a liquid, like the ocean or a lake, when the water molecules have enough energy to escape the liquid and enter the atmosphere as water vapor.

Now, let's take a look at evapotranspiration. Evapotranspiration is the combination of evaporation and transpiration. Transpiration is the process by which water is taken up by plants and released into the atmosphere as water vapor through small openings called stomata.

教師の立場で、  
入力画像に基づいて  
科学の概念を説明

In summary, evaporation is the process by which water is changed from its liquid state to its gaseous state, while evapotranspiration is the combination of evaporation and transpiration. Both these processes contribute to the water cycle and are important for maintaining the balance of water in our environment.



## Sec. 7.3 Wechsler Adult Intelligence Scale

### Prompt:

Which one of the five is least like the other four?

- Carrot
- Potato
- Corn
- Apple
- Bean

### GPT-4V:

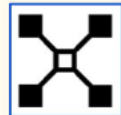
d. Apple

Explanation: Carrot, potato, corn, and bean are all vegetables, while apple is a fruit.

### Prompt:

Which one of the five is least like the other four?

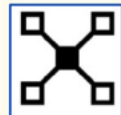
a.



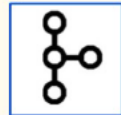
b.



c.



d.



e.



### GPT-4V:

Answer: d. is the least like the other four because it is the only one that does not have a central point connecting all the other points.

### Prompt:

Tom, twelve years old, is three times as old as his sister. How old will Tom be when he is twice as old as his sister?

- 18
- 21
- 16
- 20
- 15

### GPT-4V:

If Tom is three times as old as his sister, then his sister is one-third of his age. So, his sister is 4 years old. Now, we need to figure out when Tom will be twice as old as his sister. This means that his sister's age will be half of his age. So, we need to find a number that is double the age of 4. The answer is 8. So, Tom will be 16 years old when he is twice as old as his sister. The correct answer is c. 16.

### Prompt:

Which one of the five designs makes the best comparison?



a.



b.



c.



d.



e.



### GPT-4V:

Answer: c. The black dress shoe

Explanation: The analogy is comparing a hand to a glove and a foot to a shoe. The best comparison would be the black dress shoe as it is an item of clothing that is worn on the foot, just like a glove is worn on the hand.

## Sec. 7.4 Raven's Progressive Matrices

### RPM: Entire Question Page

### Prompt:

Find out the pattern in the first column and use it to infer the missing figure in the second column.



5. Choose a drawing that fits the empty space:



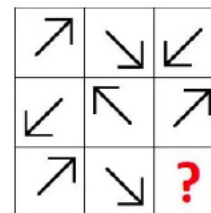
### GPT-4V:

The correct drawing that fits the empty space is the upside-down triangle. This is because the pattern in the grid is that each row and column has one of each shape in different orientations.

### Prompt:

Find out the pattern in the first two rows and use it to infer the missing figure in the third row.

Figure out which figure of A to H that should logically replace the missing figure/question-mark.



☐ A



☐ B



☐ C



☐ D



☐ E



☐ F



☐ G



☐ H

### GPT-4V:

The pattern in the first two rows is that the figures are rotated 90 degrees clockwise in each subsequent cell. Therefore, the missing figure in the third row should be a rotated version of the figure in the previous cell. Looking at the options, the correct answer is OC, as it is the figure that is rotated 90 degrees clockwise from the figure in the previous cell.

IQテスト



# 画像生成AI DALL-E3 (Open AI)

“Cinematic shot of a tiger and an young zookeeper stargazing on a boat in the middle of nowhere during night when a whale made up of stars swims nearby”



# 2024年 AIがノーベル物理学賞!!



NOBELPRISET I FYSIK 2024  
THE NOBEL PRIZE IN PHYSICS 2024



**John J. Hopfield**

Princeton University, NJ, USA



**Geoffrey E. Hinton**

University of Toronto, Canada

*"för grundläggande upptäckter och uppfinningar som möjliggör maskininlärning med artificiella neuronnätverk"*

*"for foundational discoveries and inventions that enable machine learning with artificial neural networks"*

THE  
NOBEL  
PRIZE

# **AI for Physics**

## **シミュレーションベース機械学習**

# 問題意識

## ① モデルベースアプローチ（従来のサイエンス研究）

第一原理計算によるシミュレーション(演繹的推論)

## ② データ駆動アプローチ（現状の機械学習研究）

膨大なデータからの帰納的学習

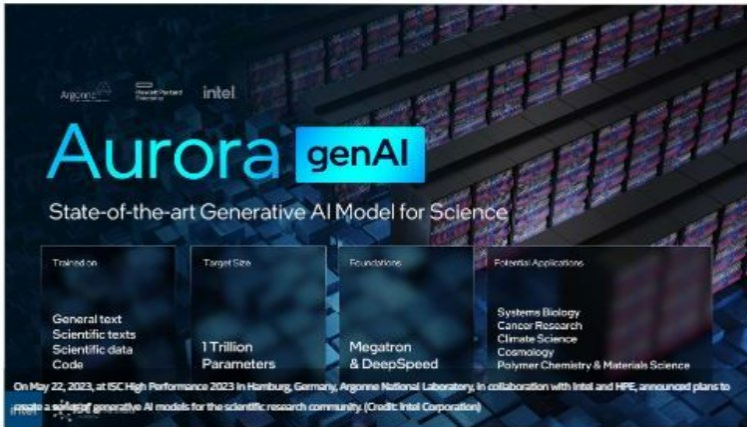
(レッドオーシャン！ ; Googleの圧勝)



いずれにおいても、**限られたデータの壁**、**内挿的予測の限界**を打破できていない！（現状ではブルーオーシャン！）



# AI for Science (世界の動向)



[https://ledge.ai/articles/intel\\_aurora-genai](https://ledge.ai/articles/intel_aurora-genai)

**Intel: 1兆パラメータの科学指向AI「Aurora gen AI」**



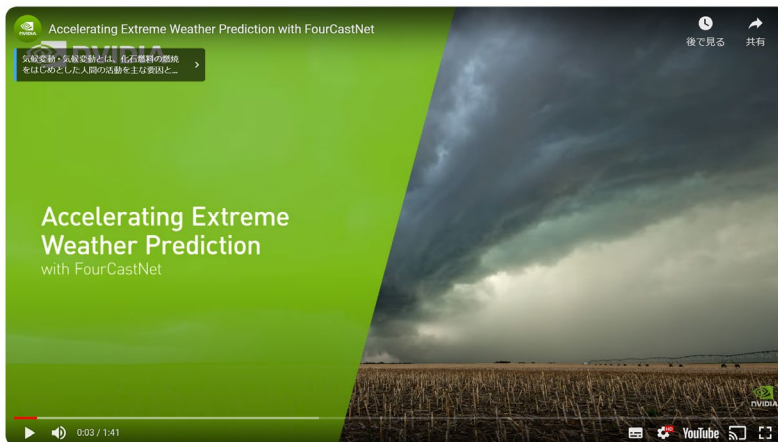
**Microsoft**

<https://www.microsoft.com/en-us/research/blog/ai4science-to-empower-the-fifth-paradigm-of-scientific-discovery/>

Accelerating Extreme Weather Prediction with FourCastNet

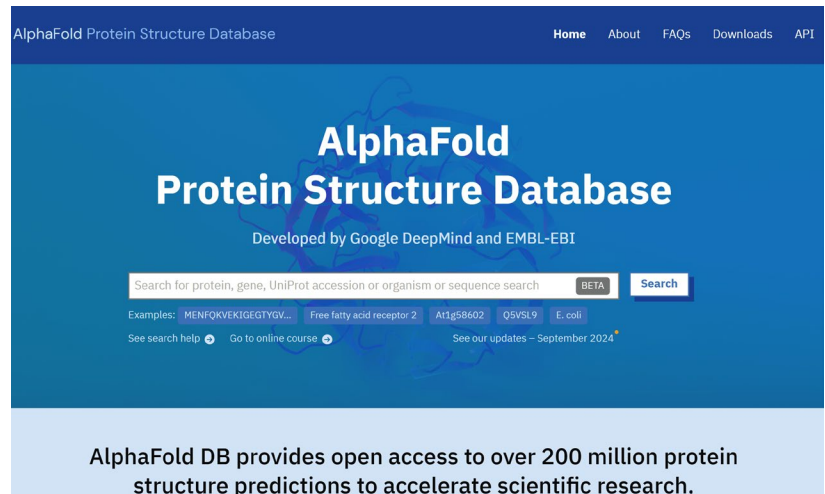
YouTube · NVIDIA · 2022/03/23

YouTube



[https://www.youtube.com/watch?v=nuT\\_U1AQz3g](https://www.youtube.com/watch?v=nuT_U1AQz3g)

**NVIDIA: 気象予測「FourCastNet」**



<https://alphafold.ebi.ac.uk/>

**Google DeepMind & EMBL-EBI: タンパク質構造予測「AlphaFold」**

# AI for Science at RIKEN

## TRIP: Transformative Research Innovation Platform of RIKEN Platforms



最先端研究プラットフォーム連携（TRIP）事業本部は、理研の最先端研究プラットフォーム群（スーパーコンピュータ「富岳」、量子コンピュータ、大型放射光施設SPring-8、バイオリソース事業など）をつなぎ、「良質なデータ整備」「AI×数理」「計算可能領域の拡張」による研究DXの先駆的な研究基盤を機能させるとともに、これらを活用した新たな価値創生に資する研究を推進します。

TOP ABOUT TRIP RESEARCH NEWS

MESSAGE CONTACT

ENGLISH

### TRIP 1

良質なデータ整備  
(蓄積・統合)

拡張した計算基盤による解析に向けた良質なデータ取得と多様なデータの蓄積・統合による解析基盤の整備

READ MORE

### TRIP 2

AI×数理  
(予測の科学)

数理科学により、スパコン、AI、量子コンピュータをつなぎ、様々な分野における量子HPCハイブリッド計算のアルゴリズム開発を実施

READ MORE

### TRIP 3

計算可能領域の拡張  
(量子HPCハイブリッド)

量子コンピュータ・スーパーコンピュータを連携させたハイブリッドコンピューティング基盤による計算可能領域の拡張

READ MORE

### Use case

新たな価値創生、社会課題解決への貢献

プラットフォームオププラットフォームを活用したユースケースによる新たな価値創生、社会課題解決への貢献

READ MORE

<https://trip.riken.jp/>

# ～モデル駆動とデータ駆動の新たな融合～

## ～ AI for Science ～

- シミュレーションを深層学習で代替
- 物理法則に忠実な深層学習の実現
- 複雑な数理モデルを深層学習で解く

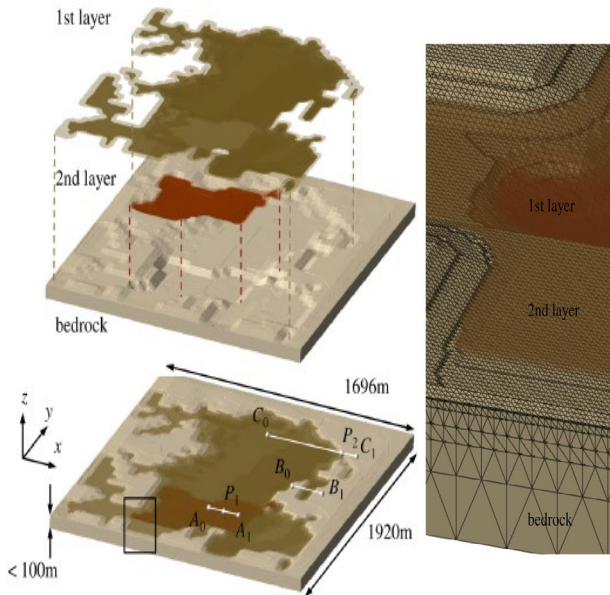


# シミュレーションを 深層学習で代替

シミュレーションの高速化

# 地震動予測

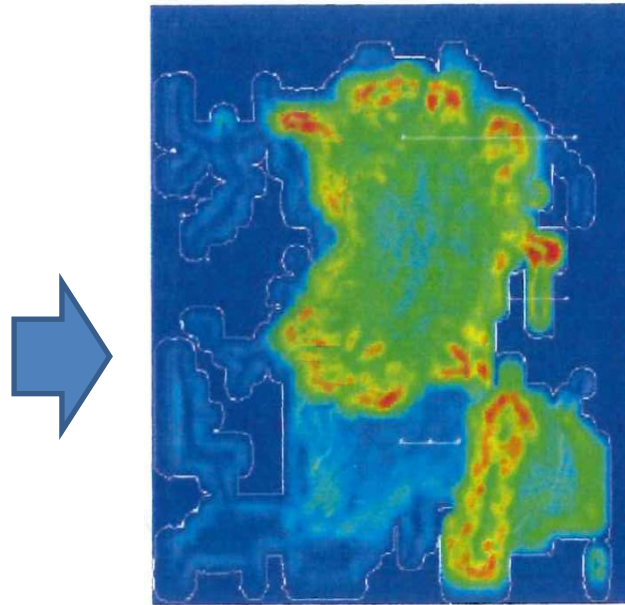
## 3D地盤モデル



### 3D分布

- $V_P$ : P-wave speed
- $V_S$ : S-wave speed
- $\rho$ : mass density

## 地震動シミュレーション



FEM IIによる大規模シミュレーション

$$V_P^2 \Delta \phi \quad \phi = \nabla \cdot \mathbf{u}$$

$$\ddot{\psi} = V_S^2 \Delta \psi \quad \psi = \nabla \times \mathbf{u}$$

## 地震被害推定



建物、インフラ被害  
予測

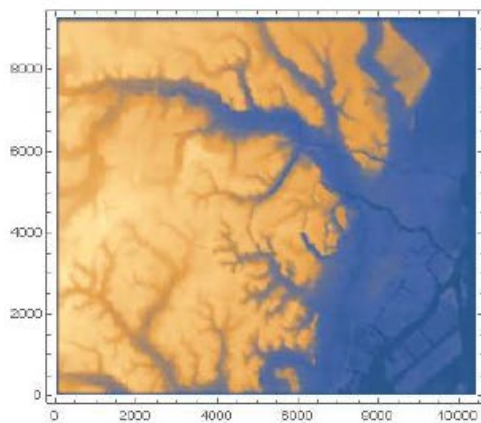
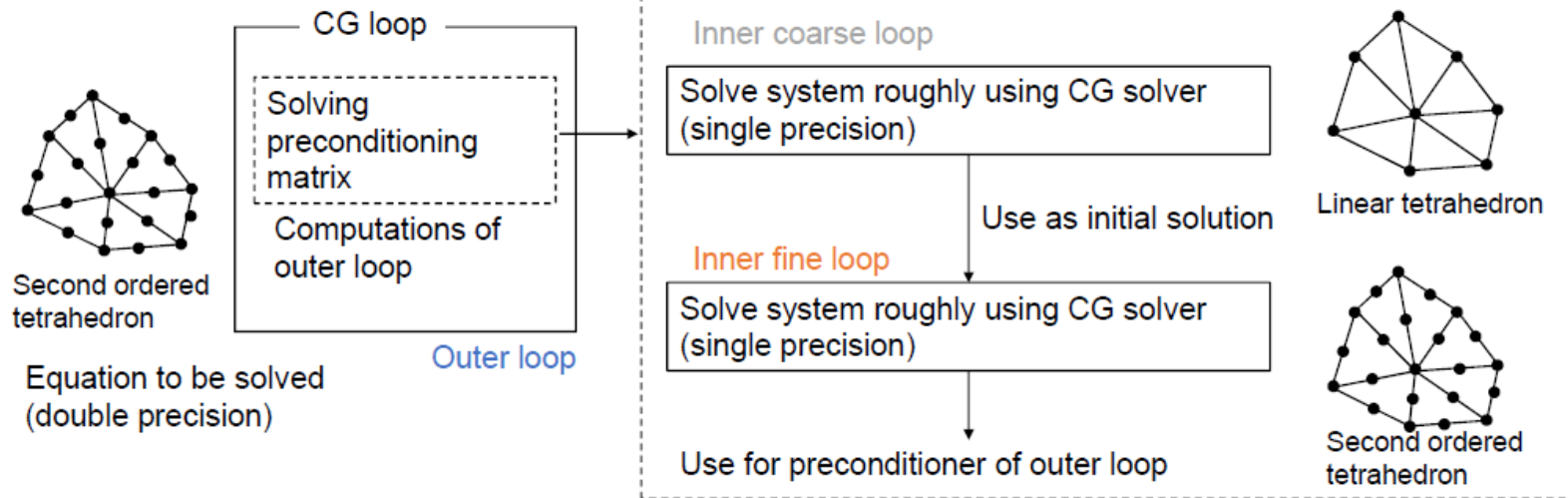
メッシュ間隔  
<100 m

# HPC上での物理シミュレーション

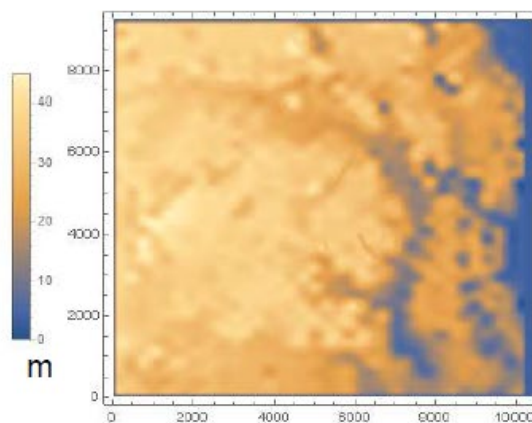
## HPC手法の開発

SC15 ゴードンベル賞ファイナリスト<sup>ref6)</sup>

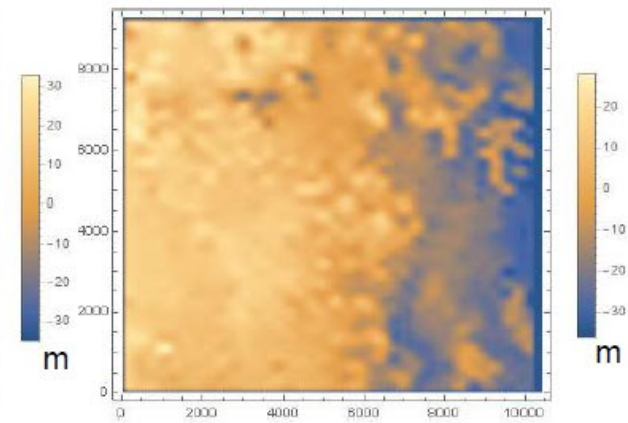
市村(東大)



地表面標高

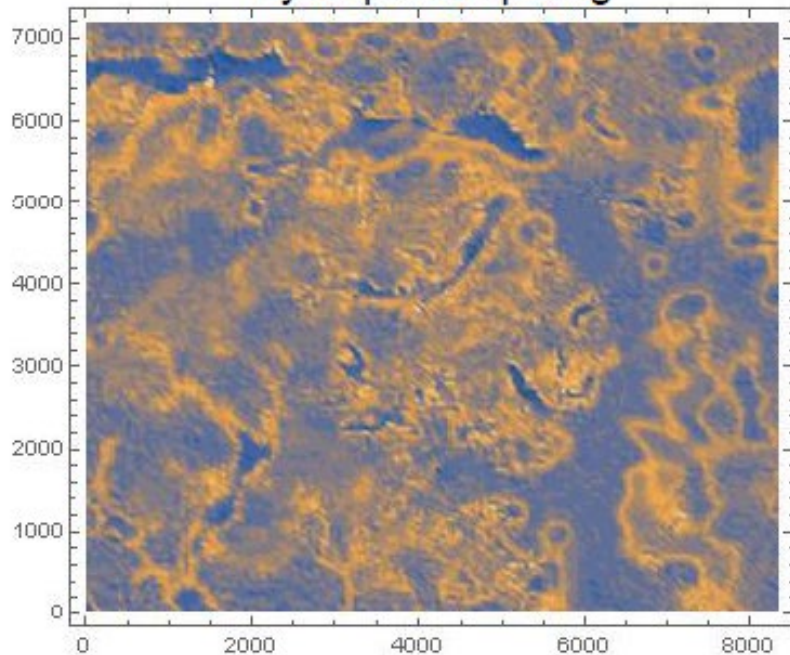


1層目と2層目の層境標高

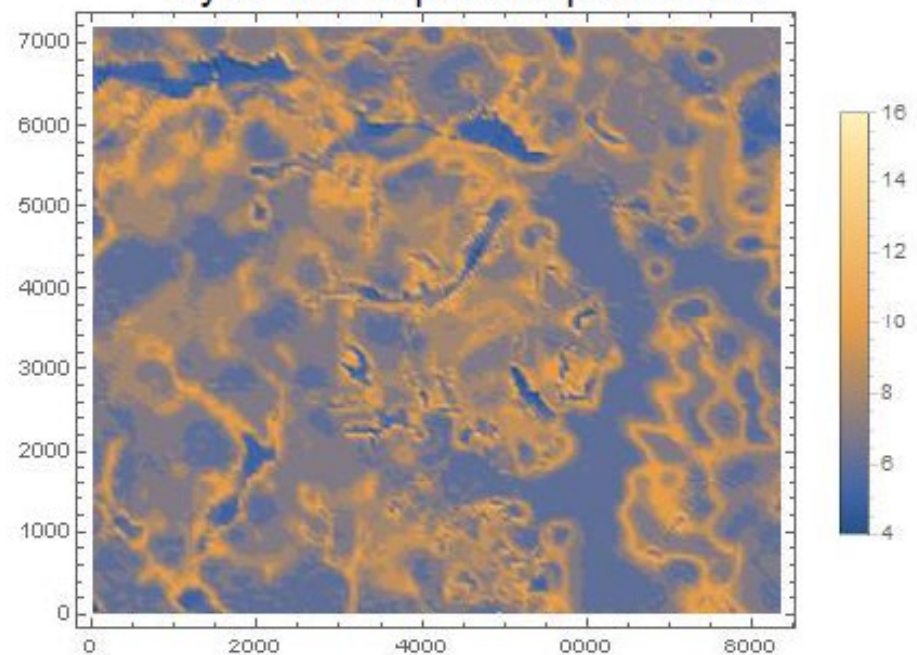


2層目と3層目の層境標高

By Supercomputing



By AI with Super-computed data



	Training data	Test data	Check data	Total
# of samples	22360	1000	1000	24360
Case A	19929 (89.1%)	880 (88.0%)	891 (89.1%)	21700 (89.1%)
Case B	22219 (99.4%)	992 (99.2%)	994 (99.4%)	24205 (99.4%)

**SC2017  
Best Poster  
Award**

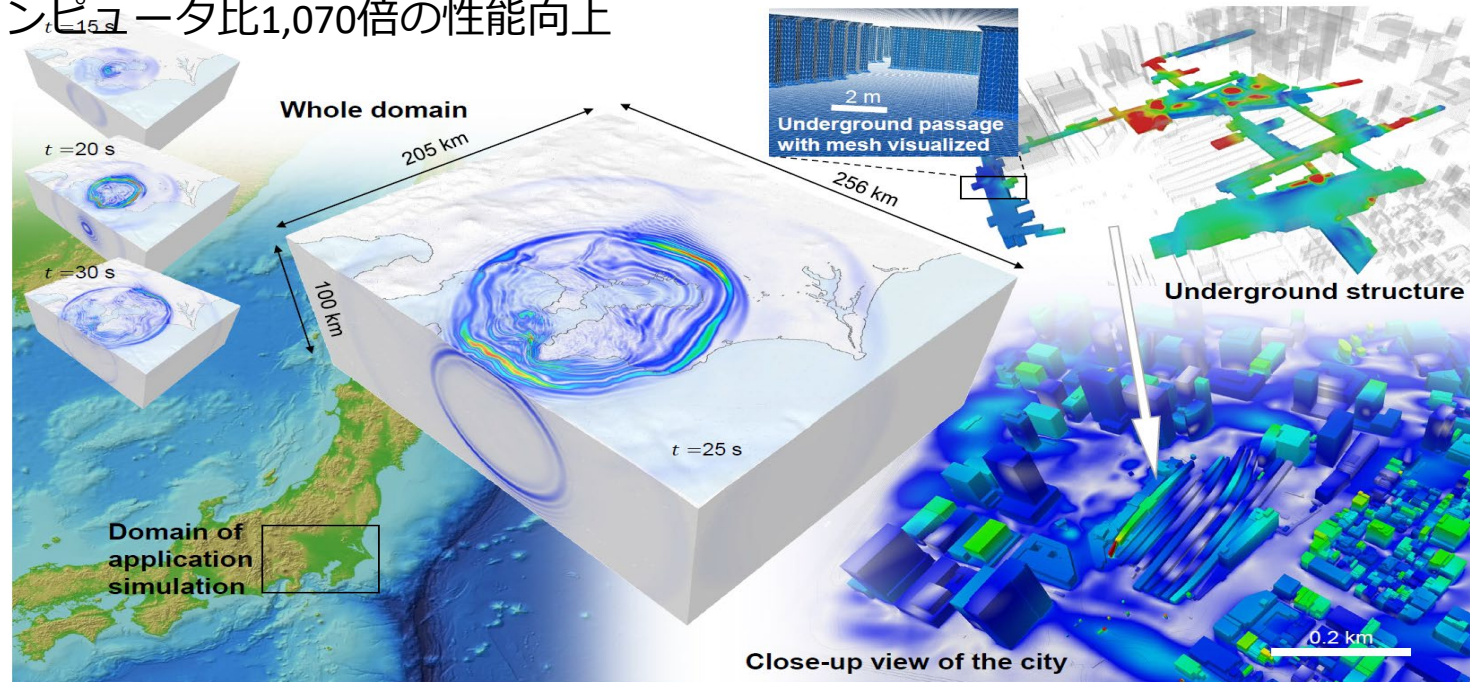
The number of correct answers and correct ratios. Judged correct if  $|\text{target value} - \text{estimated value}| < 2$   
1000 points are picked up from 23,460 points as test data and check data.



# 超並列計算物理シミュレーションと データ学習のハイブリッド手法

富岳全系（7,312,896並列）までスケールする超並列計算物理シミュレーションとデータ学習のハイブリッド手法を開発し、**世界初の断層-都市超高詳細解析を実現**

- ・地震シミュレーションの根幹となる非構造低次有限要素法によるテラ自由度級大規模非線形動的解析。地震の被害想定・メカニズム解明への貢献が期待される
- ・富岳全系までスケールするデータ学習により物理シミュレーションを高速化する手法を開発
- ・この規模・性能の地震シミュレーションは日本でのみ実現。
- ・京コンピュータ比1,070倍の性能向上



# 地震による津波浸水被害のリアルタイム予測

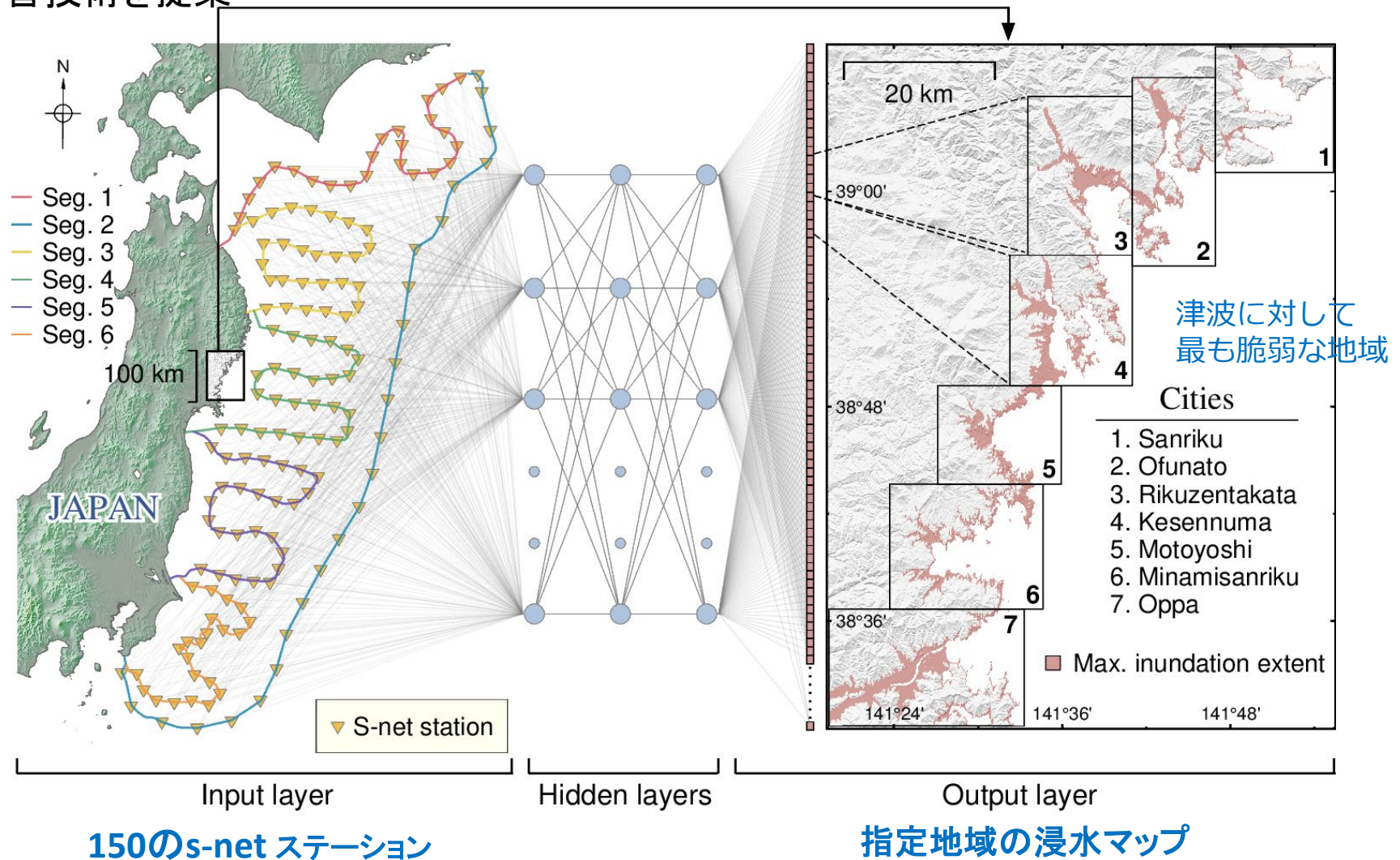
## **Machine Learning-based Tsunami Inundation Prediction Derived from Offshore Observations**

Mulia, I, Ueda, N., Miyoshi, T., Gusman, A., Satake, K.

**Nature Communications. 2022.**

# 提案アプローチ

海域の観測無しに、S-net を入力として、指定地域の浸水マップを出力する機械学習技術を提案



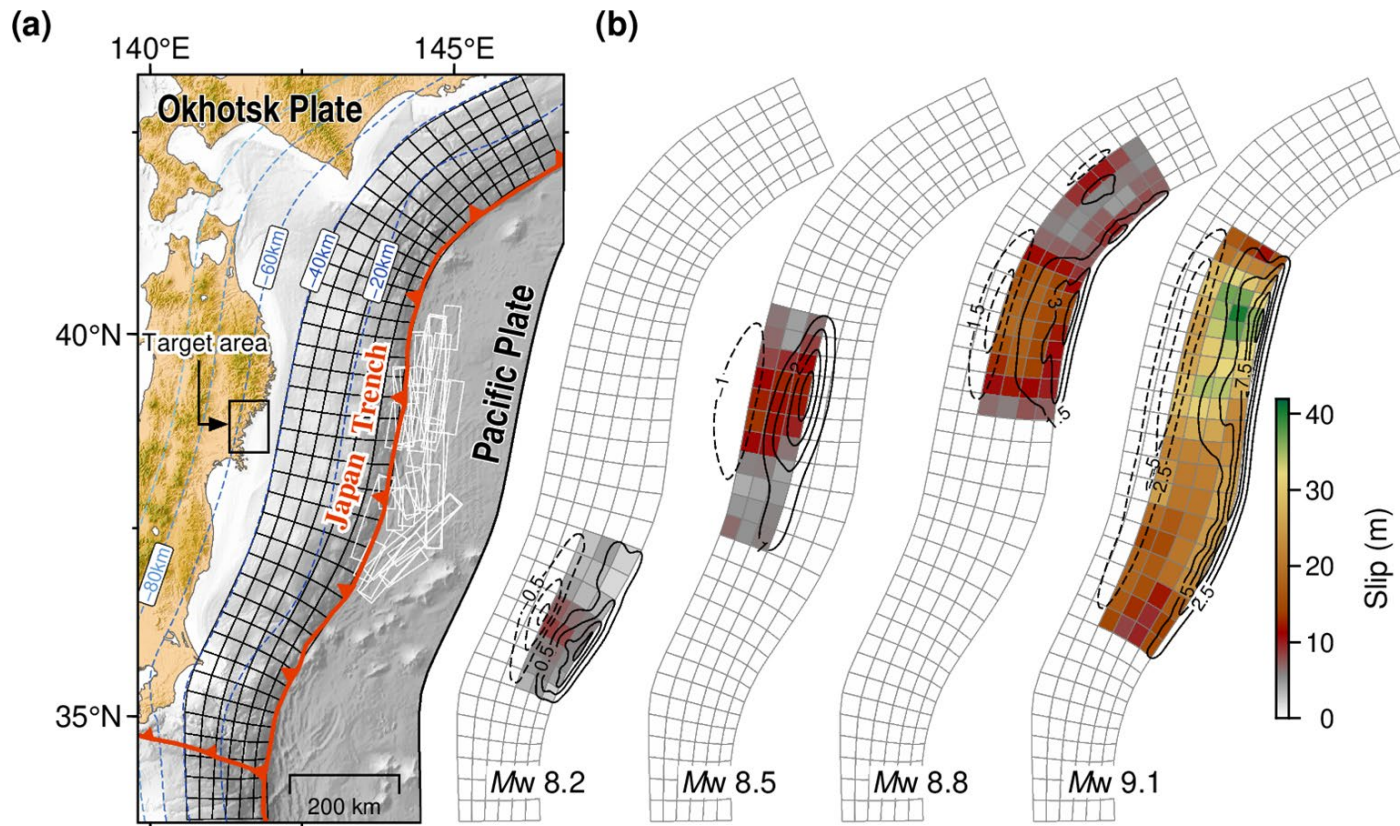
S-net: 日本海溝に沿って地震と津波のための海底観測ネットワーク (S-net)。

1 kHz の高周波データをリアルタイムで送信できる海底圧力センサーと地震計で構成



# 地震発生シナリオ

- $M_w$  8.0~9.1 (0.1間隔) の仮想のメガスラスト地震を多数生成し、スリップを補間して曲面パッチに投影する。
- 非線形浅水波方程式を数値的に解くことにより津波の伝搬を算出。
- 曲面パッチ(グリッド)での値(最大水面高度)をS-netのセグメントの値に変換。



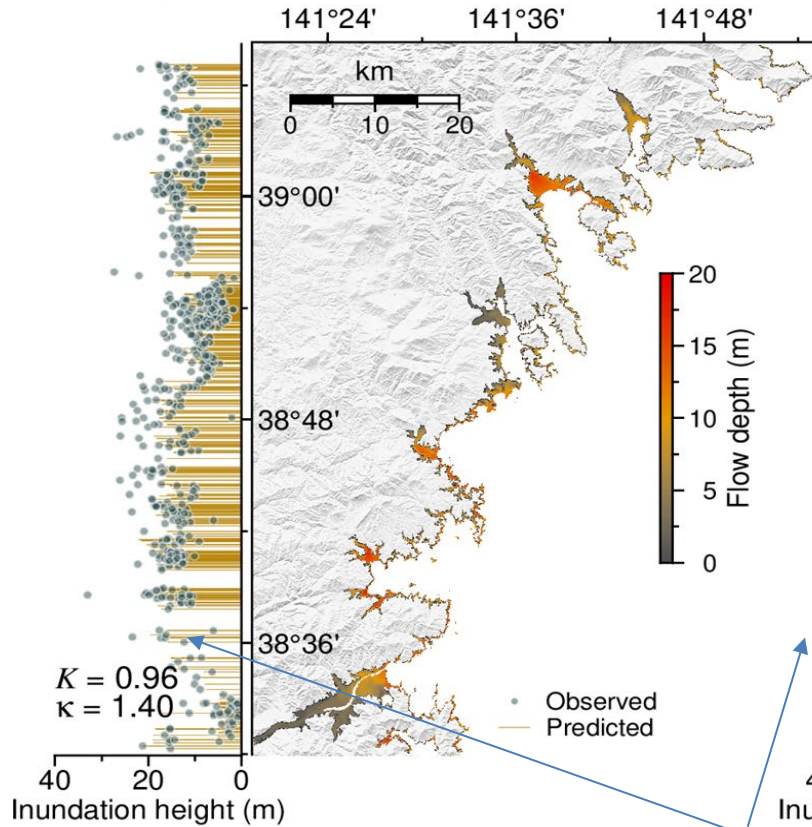
メガスラスト断層(地震の沈下と隆起を表す)



# 検証実験

## 物理シミュレーション

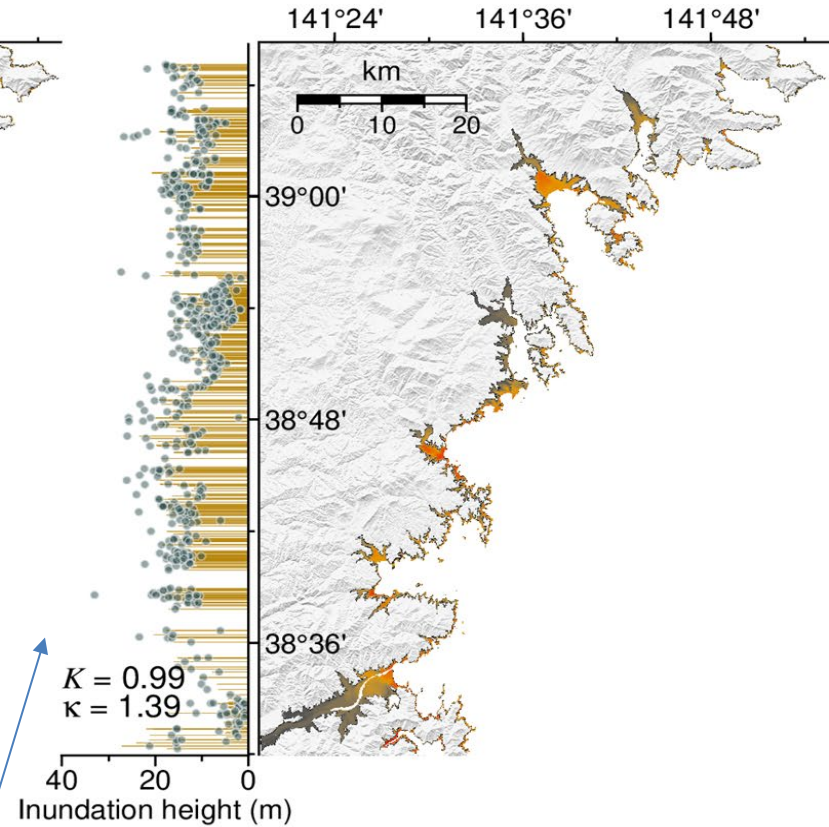
(a) Physics-based model



Computing time = ~30 min

## 提案法

(b) Machine learning-based model



Computing time = 0.05 sec

同等な予測精度をリアルタイムで実現！

# 物理法則に忠実な 深層学習の実現

～ AI for Science ～

# 研究動向

## DNNでODEを解く

ベクトル場 $f(x;\theta)$ を  
直接NN等で学習

- **Neural ODE** [Chen+'18]
  - DiffGP [Hegde+'19]
- など

## 偏微分方程式を NNで解く

- PDEと初期・境界条件  
を与えてNNで学習
- **PINN**[Raiss+'19]など

## 物理法則を満たしつつ データからモデルを獲得

保存量（エネルギー等）  
という構造を仮定  
（保存系）

- **HNN** [Greydanus+'19]
  - LNN [Cranmer+'20]
  - SympGPR [Rath+'21]
- など

エネルギーの減衰という  
知識を仮定  
（散逸系）

- **DGNet** [Matsubara+'20]
  - GHNN [Course+'20]
- など

## データから解作用素を学習

初期条件とその場合のPDEの解の組  
のデータから背後にあるPDEを推定  
**Neural Operator**[Li+'20]など

# Neural Ordinary Differential Equation (NODE)

Residual NN (**RESNet**: He et al., NeurIPS2015)

Chen et al., NeurIPS2018

**skip connection の導入**

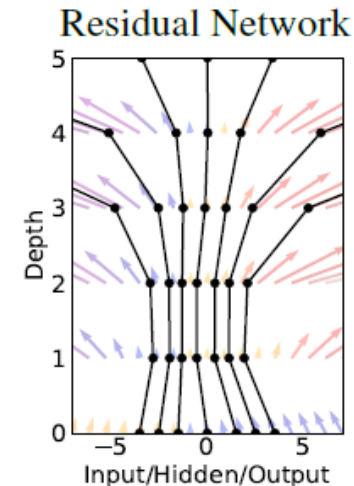
Traditional Deep Nets

```
h1 = f1(x)
h2 = f2(h1)
h3 = f3(h2)
h4 = f3(h3)
y = f5(h4)
```

ResNets

```
h1 = f1(x) + x
h2 = f2(h1) + h1
h3 = f3(h2) + h2
h4 = f4(h3) + h3
y = f5(h4) + h4
```

$$h_N = f(\theta, h_{N-1}) + h_{N-1}$$



**連続(時間)隠れ層への拡張: NODE**

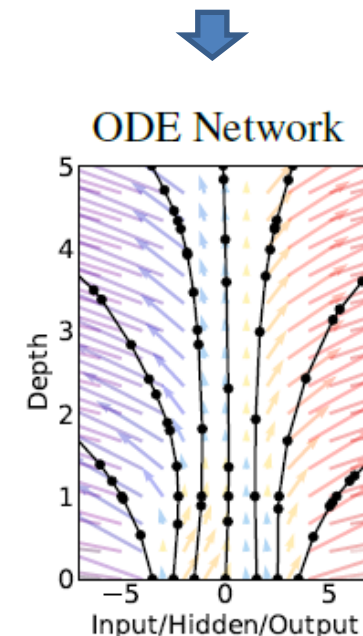
**共通の関数  $f$  をDNNで代替:**  $\frac{dh(t)}{dt} = f(h(t), t, \theta)$

ODEの解 $z(t_1)$ を入力とした損失関数

$$\begin{aligned} L(z(t_1)) &= L\left(z(t_0) + \int_{t_0}^{t_1} f(z(t), t, \theta) dt\right) \\ &= L(\text{ODESolve}(z(t_0), f, t_0, t_1, \theta)) \end{aligned}$$

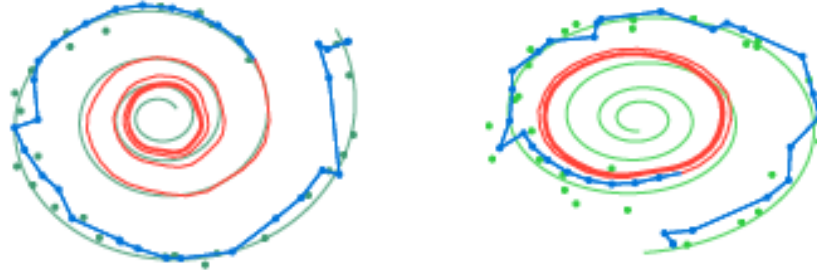
NNの順伝搬+オイラー法

Lの $\theta$ に関する微分: アドジョイント法によりメモリーの効率化を図っている

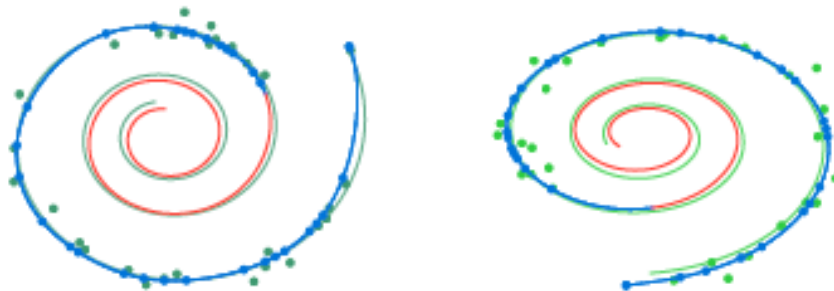




# 時系列データの学習例



(a) Recurrent Neural Network



(b) Latent Neural Ordinary Differential Equation

緑の点が時系列データを欠損あり&不等間隔にした観測値、青の線がモデルがフィットした線、赤の線が予測値

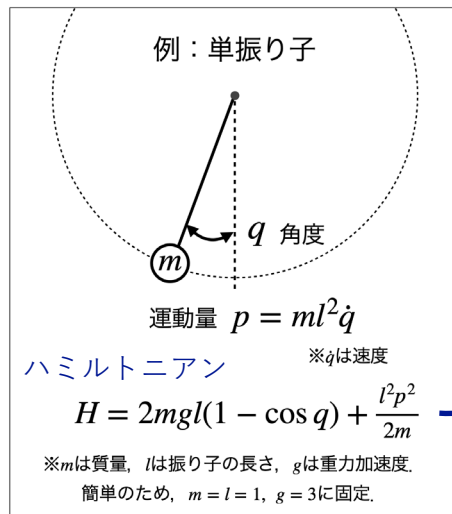
Chen et al., NeurIPS2018

# ハミルトニアンNN

ハミルトン力学系を仮定し、エネルギー保存のような「保存則」を満たすようにDNNを学習 [1,2]

## ハミルトン力学とは

位置 $q$ と運動量 $p$ で決まる関数 $H$ （ハミルトニアン）が与えられると、ハミルトンの運動方程式を通じてダイナミクスを知ることができる



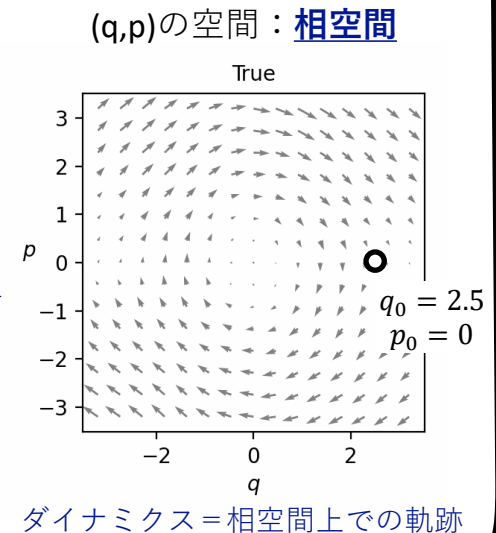
ハミルトンの  
運動方程式

$$\dot{q} = \frac{\partial H}{\partial p}$$
$$\dot{p} = -\frac{\partial H}{\partial q}$$

$(q,p)$ の動きを決める  
ベクトル場

+

$(q_0, p_0)$   
初期値



[1]Samuel Greydanus et al., Hamiltonian Neural Networks, NeurIPS, 2019.

[2]Takashi Matsubara et al., Deep Energy-Based Modeling of Discrete-Time Physics, NeurIPS, 2020.

# 複雑な数理モデルを 深層学習で解く

～ 数値(離散)解法から連続解法へ ～

# 物理深層学習による地殻変動解析

## Physics-Informed Deep Learning Approach for Modeling Crustal Deformation

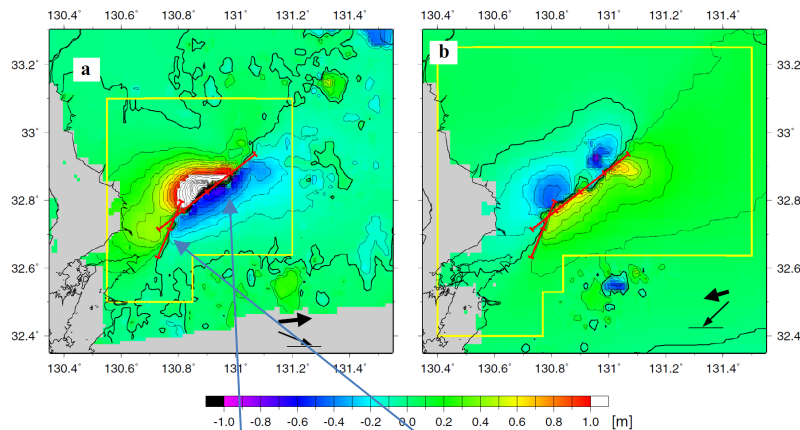
Okazaki, T., Hirahara, T, Ito, T, Ueda, N.

Nature Communications. 2022.



# 地殻変動のモデリング

観測(人工衛星による地表変位データ)



地表面の変形  
(観測)

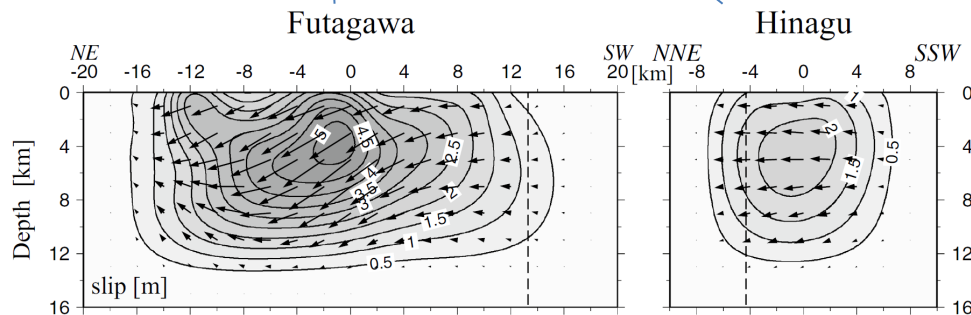
例. 2016年熊本地震

地殻ダイナミクスの理解

- 地震断層の破壊様式
- プレーートの運動様式
- 地下の弾性・粘性構造

順問題  
本研究

逆問題



地下の断層・プレート運動  
(モデル)

半世紀以上に亘り、弾性体力学に基づく  
解析的・数値的計算手法が発展してきた

本研究 PINNによるモデリングにより変動や  
任意の地形・地下構造を連続的に  
(離散化せずに) 表現可能

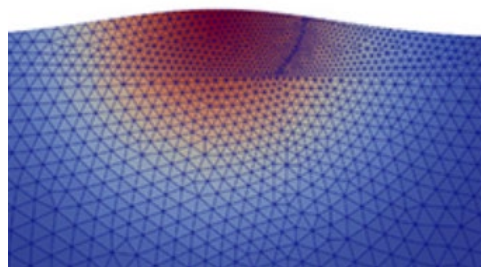
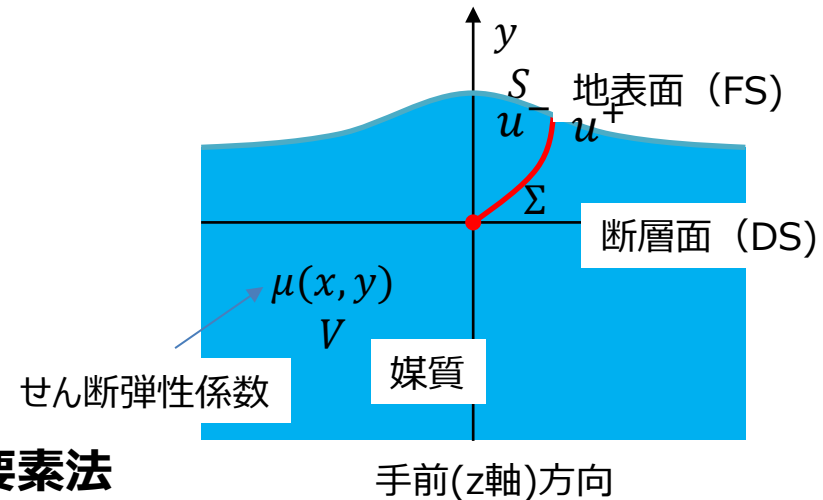
# 地殻変動をシミュレートする機械学習技術

## ■地殻ダイナミクス(非線形微分方程式)を制約条件とする深層学習技術を考案

例) 布田川断層(熊本地震, 2016)

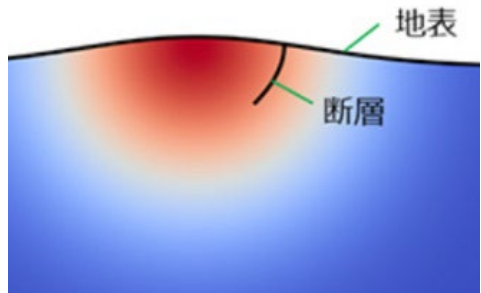


視線(z軸)方向



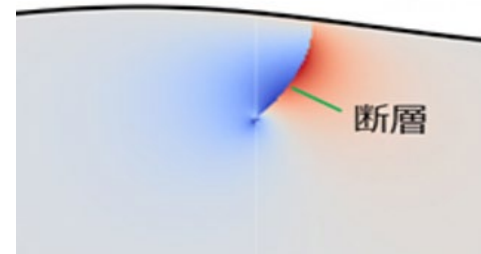
有限要素法  
(従来法)

メッシュ毎の  
線形近似



提案法

メッシュフリーな  
厳密解析



従来手法(上)と本手法(下)における解析領域の表現方法。図は地下構造の断面図。配色は地下を構成する岩石の硬さを表わす。

本手法による地殻変動の解析結果。

地下構造の断面図。配色は地震による地殻の変形の計算結果を表す。断層や地形の連続的に曲がった形状や、断層や地下構造の急激な変化によるひずみが解析できている。

# PINNによる横ずれ断層の地殻変動解析

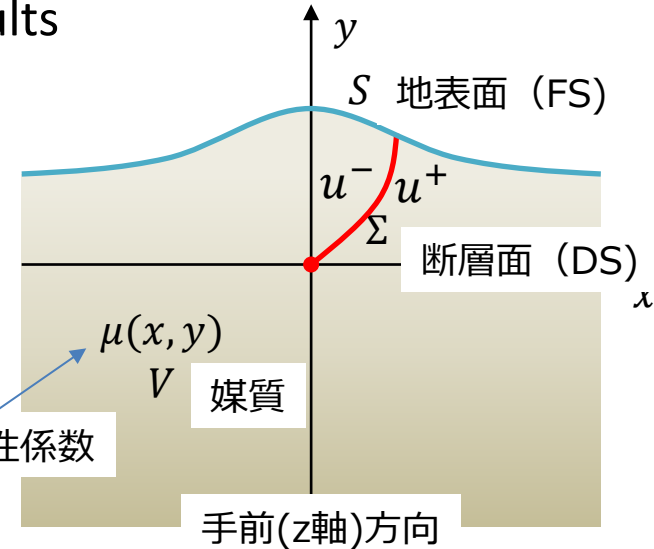
Strike-slip faults

例) 布田川断層(熊本地震, 2016)



視線(z軸)方向

せん断弾性係数



**支配方程式** (横ずれ断層:  $u(x, y)$  は  $z$  軸方向で一定)

DS: dislocation surface  
FS: free surface

$$\mu \nabla^2 u + \nabla \mu \cdot \nabla u = 0 \text{ in } V \quad \leftarrow v \text{ における微分方程式}$$

$$u^+ - u^- = s \text{ on } \Sigma \quad \leftarrow \Sigma \text{ における境界条件 (s: 既知の変位距離)}$$

$$\boldsymbol{\sigma}^+ \cdot \mathbf{n}^{\text{DS}} = \boldsymbol{\sigma}^- \cdot \mathbf{n}^{\text{DS}} \text{ on } \Sigma \quad \leftarrow \Sigma \text{ における境界条件} \quad \mathbf{n}^{\text{DS}} = (n_x^{\text{DS}}, n_y^{\text{DS}}, 0)$$

$$\boldsymbol{\sigma} \cdot \mathbf{n}^{\text{FS}} = 0 \text{ on } S \quad \leftarrow S \text{ における境界条件} \quad \mathbf{n}^{\text{FS}} = (n_x^{\text{FS}}, n_y^{\text{FS}}, 0)$$

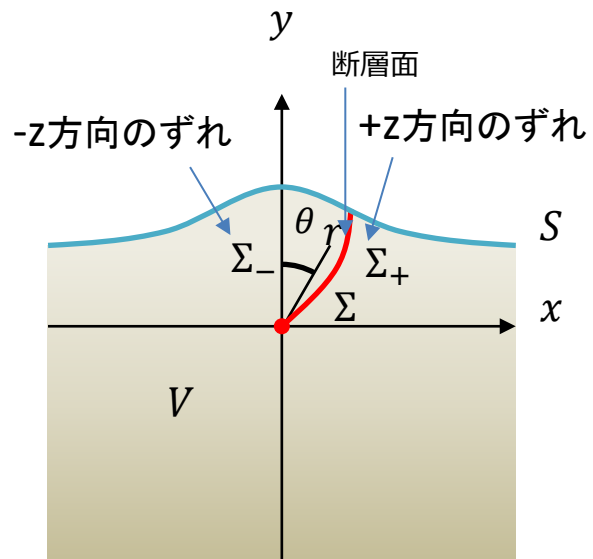
$\boldsymbol{\sigma}$ : ストレステンソル (3x3 行列)

Note: 地殻変動解析では定常状態での解析故、空間微分のための微分方程式となる

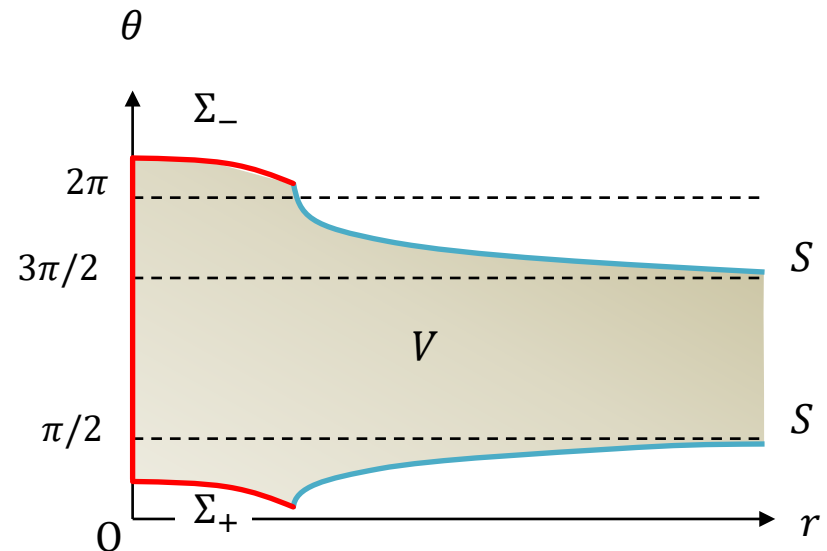
# PINNの適用における課題と解決法

地震により断層が破壊される(ずれ動く)ため、断層の両側( $\Sigma_+$ ,  $\Sigma_-$ )で $u(x, y)$ が不連続となり、NNで用いる連続関数では表現不可能

➡ 断層面 $\Sigma$ に沿ってbranch cutをもつ**極座標表示**を導入



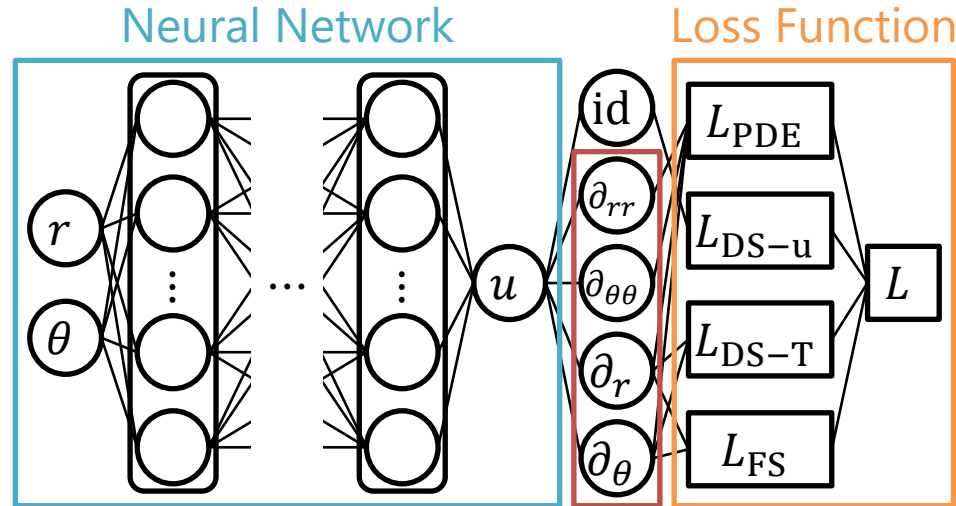
断層の両側で $u(x, y)$ の値が不連続となりPINNの適用は不可



$u(r, \theta)$ は領域全体で連続関数  
変位不連続を $\Sigma_+$ と $\Sigma_-$ の境界条件として表現



# PINNによる定式化



実験では、8層、  
40ノード/層のNN

Automatic Differentiation

## 損失関数：

$$L = L_{\text{PDE}} + L_{\text{DS-u}} + L_{\text{DS-T}} + L_{\text{FS}} \quad \leftarrow \text{微分方程式と境界条件に基づく損失関数}$$

$$L_{\text{PDE}} = [r^2 u_{rr} + r u_r + u_{\theta\theta} + \mu^{-1}(r^2 \mu_r u_r + \mu_\theta u_\theta)]^2 \quad \leftarrow \text{Vにおける微分方程式}$$

$$L_{\text{DS-u}} = (u^+ - u^- - s)^2 \quad \leftarrow \Sigma \text{における境界条件 (s: 既知の変位距離)}$$

$$L_{\text{DS-T}} = r^2 [\mu^+ (n_x^{\text{DS}} u_x^+ + n_y^{\text{DS}} u_y^+) - \mu^- (n_x^{\text{DS}} u_x^- + n_y^{\text{DS}} u_y^-)]^2 \quad \leftarrow \Sigma \text{における境界条件}$$

$$L_{\text{FS}} = r^2 (n_x^{\text{FS}} u_x + n_y^{\text{FS}} u_y)^2 \quad \leftarrow \text{Sにおける境界条件}$$

$$u_x = \sin \theta u_r + r^{-1} \cos \theta u_\theta$$

$$u_y = \cos \theta u_r - r^{-1} \sin \theta u_\theta$$

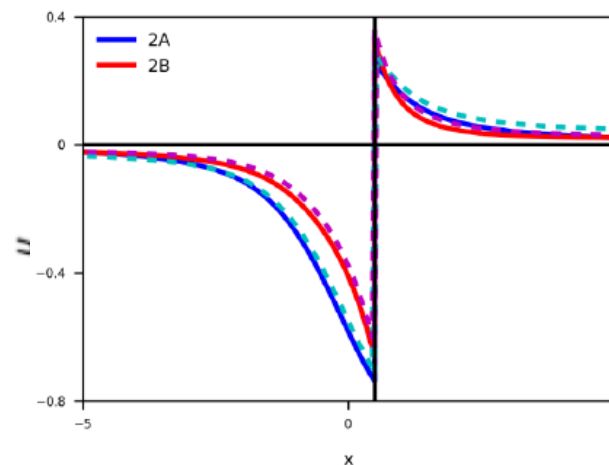
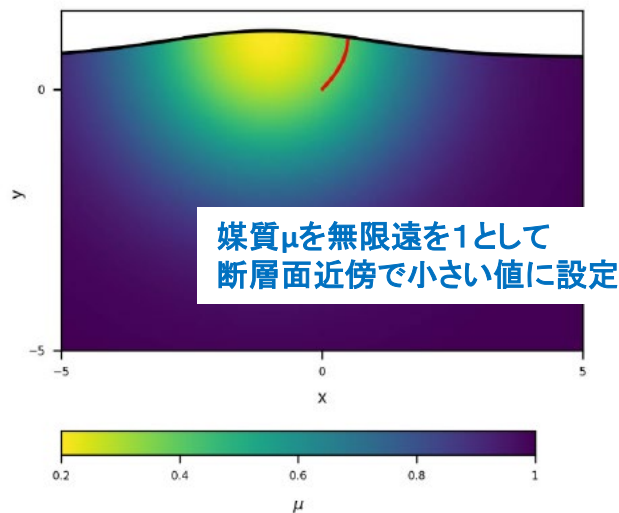
Note: 下付き添え字はその変数での微分を表す

# 実験結果（断層が曲線、地表面が曲線、媒質非一様のケース）

解析的に解けないため、  
有限要素法(FEM)と比較

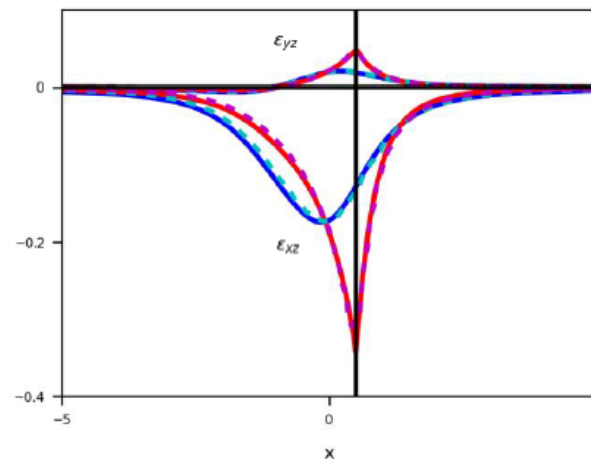
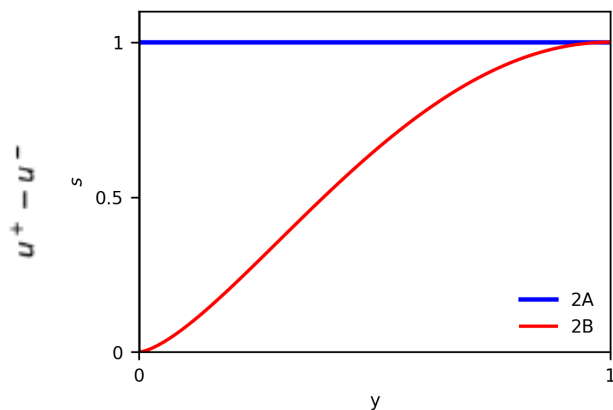
タスク

結果



点線:FEM  
実線:PINN

地表面での $u$ (変位)



地表面での歪量( $u$ の空間微分)

PINNの結果は、有限要素法(FEM)と同等

54

# 作用素学習(Operator Learning)

PDEの解作用素の学習

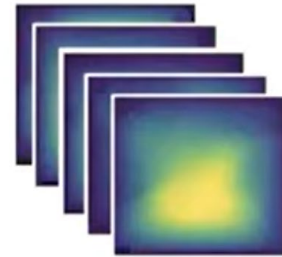
モデルパラメータ

Goal: N組の学習データ  $\{(a_j, u_j)\}_{j=1}^N$  から、解作用素  $\mathcal{G}_\theta: \mathcal{A} \rightarrow \mathcal{U}$  を学習

学習



Input:  
coefficients



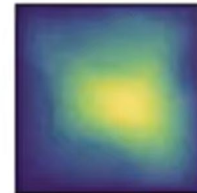
Output:  
solutions

数値シミュレーション  
により、係数と解との  
ペアのデータで学習

テスト



Input: coefficient



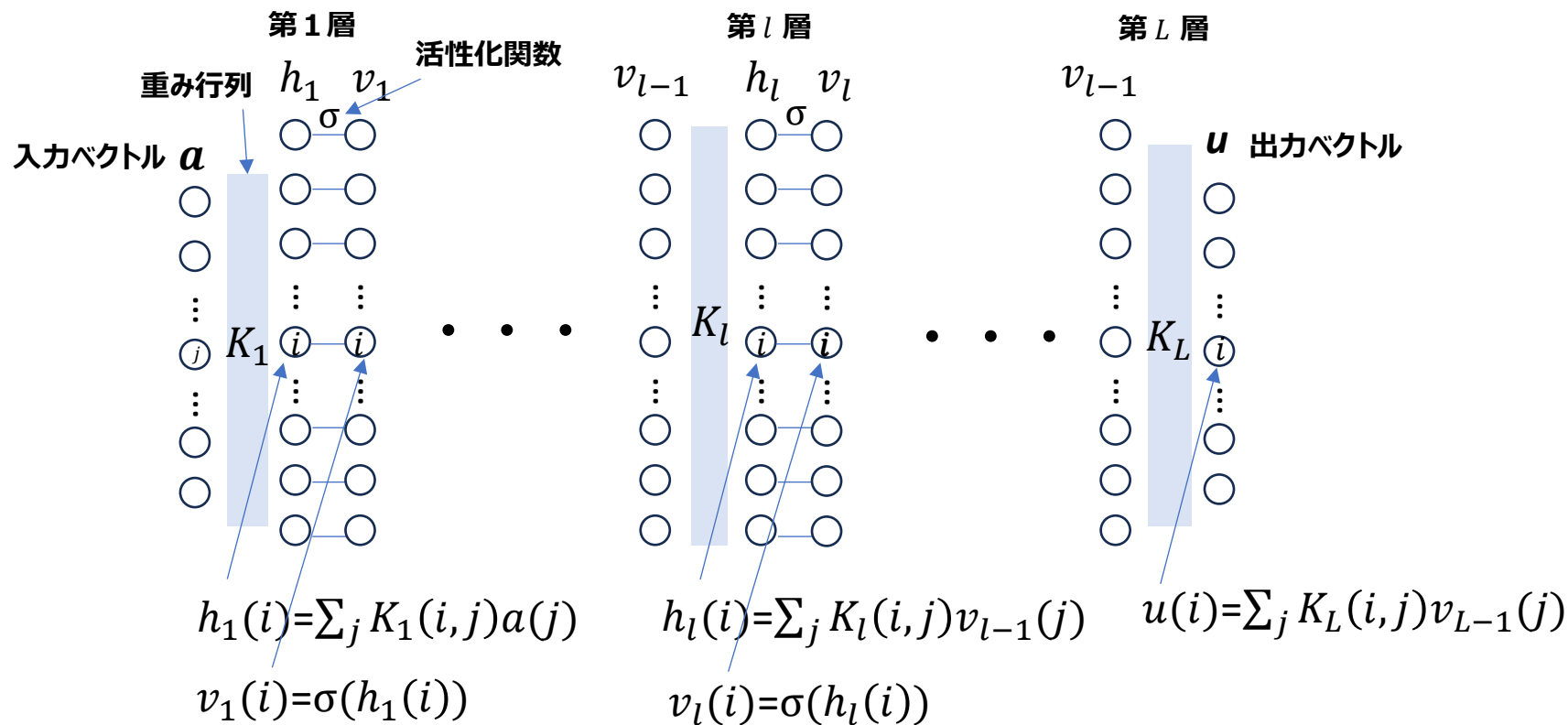
Output: solution

学習後、任意の係数  
に対する解は瞬時に  
出力可能

従来の数値解法では、異なる係数ごとに解きなおしていたが、  
作用素学習では一括学習が可能。ただし、汎化能力を上げる  
には、多数の学習データが必要

# 復習) 多層ニューラルネットワーク

通常のDNNは以下のように、積和計算(線形演算)と活性化関数(非線形演算)を繰り返している(簡単のためバイアス項は無視)



**$L$ 層DNNのベクトル、行列表記**

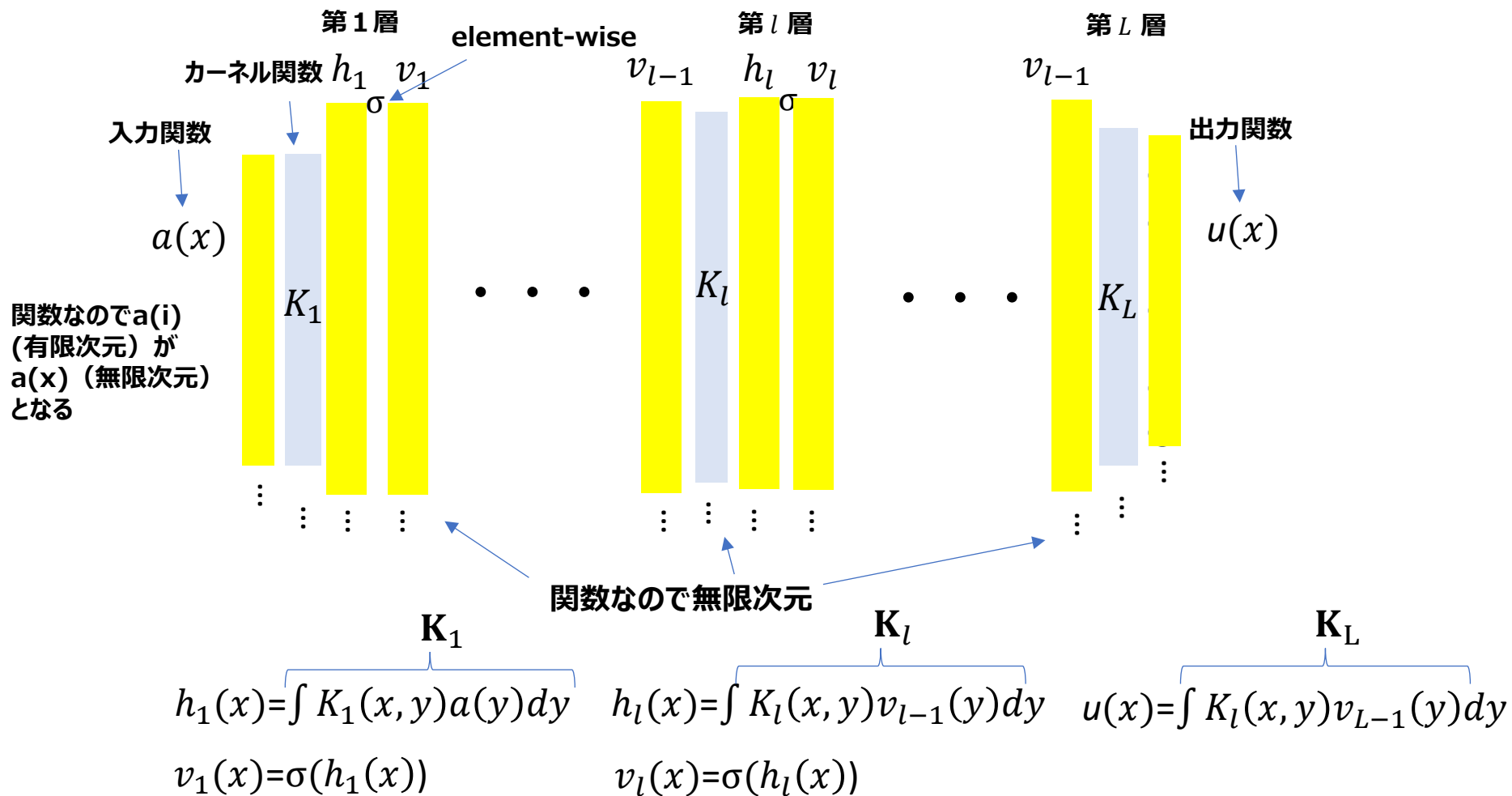
$$u = G(a) = K_L(\sigma(K_{L-1}(\sigma(\dots \sigma(K_1(a) \dots)))$$

Note: element-wiseの演算



# 無限次元拡張

関数間の写像を実現するには、NNの各層のノード(離散)を連続化したDNNを考えればよい



$$u = G(a) = K_L(\sigma(K_{L-1}(\sigma(\dots \sigma(K_1(a) \dots)))$$

**NOはDNNでのノードを連続拡張したものを見せる！**

# Neural Operator

GNO: Li+ , arXiv'20

$$\mathcal{G}_\theta := \mathcal{Q} \circ \sigma_T(W_{T-1} + \mathcal{K}_{T-1} + b_{T-1}) \circ \cdots \circ \sigma_1(W_0 + \mathcal{K}_0 + b_0) \circ \mathcal{P}$$

活性化関数

T層のカーネル積分作用素

バイアス

- $a(x)$ から $u(x)$ への変換をT層のNNでモデル化

$v_t(x)$ は $\mathbb{R}^{d_v}$ 上のrepresentation,  $P(\cdot)$ と $Q(\cdot)$ はNN

$v_0(x) = P(a(x))$  **Lifting: 次元を上げる**

For  $t=1, \dots, T$

線形変換

**Iterative Kernel Integration**

$$v_{t+1}(x) := \sigma \left( W v_t(x) + (\mathcal{K}(a; \phi) v_t)(x) \right)$$

activation

Kernel integral operator

$u(x) = Q(v_T(x))$  **Projection  $u$ の次元に戻す**

グリーン関数法との関連

$$(\mathcal{L}_a u)(x) = f(x)$$

$$u(x) = \int_D G_a(x, y) f(y) dy$$

グリーン関数G

$f(y)$ を $G(x, y)$ の重みで重ね合わせる

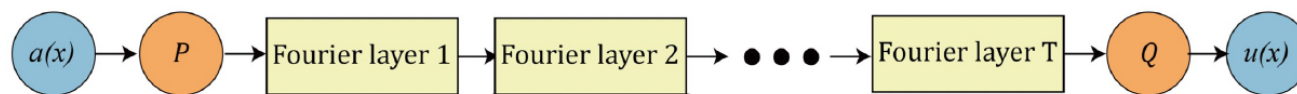
Kernel integral operator

$\kappa$ は $\mathbb{R}^{d_v \times d_v}$ の出力を持つNN

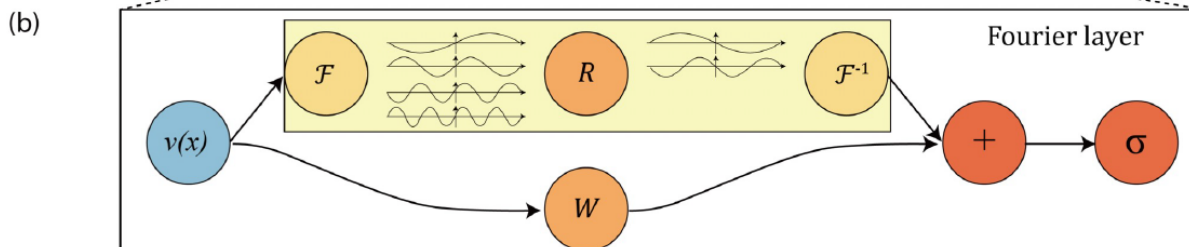
$$(\mathcal{K}(a; \phi) v_t)(x) := \int_D \kappa(x, y, a(x), a(y); \phi) v_t(y) dy \longrightarrow$$

論文では離散点をグラフと見なして  
グラフNNを用いている。Nystrom近似  
で効率化 (それでも高コスト)

# FNO: Fourier Neural Operator



U(x)の誤差を最小化  
するようにパラメータを  
学習する



$$(\mathcal{K}(a; \phi)v_t)(x) := \int_D \kappa(x, y, a(x), a(y); \phi) v_t(y) dy$$

↓ カーネル積分作用素を周波数領域での  
畳み込み積分に置き換える

$$(\mathcal{K}(a; \phi)v_t)(x) = \mathcal{F}^{-1}(\mathcal{F}(\kappa_\phi) \cdot \mathcal{F}(v_t))(x)$$

畳み込み定理：元の空間における畳み込み積分は周波数領域では積になる

$$(\mathcal{K}(\phi)v_t)(x) = \mathcal{F}^{-1}\left(R_\phi \cdot (\mathcal{F}v_t)\right)(x)$$

↑ 高周波をカット

計算量：低周波成分だけだと $O(Nk_{\max})$ , 高速フーリエ変換を用いると $O(N \log N)$

# PINO: PINNとNOの融合モデル

## Physics-Informed Neural Operator for Learning Partial Differential Equations

Zongyi Li\*, Hongkai Zheng\*, Nikola Kovachki, David Jin, Haoxuan Chen,  
Burigede Liu, Kamyar Azizzadenesheli, Anima Anandkumar

\*Both authors contributed equally to this research.

California Institute of Technology,  
Pasadena, USA

CalTech一派

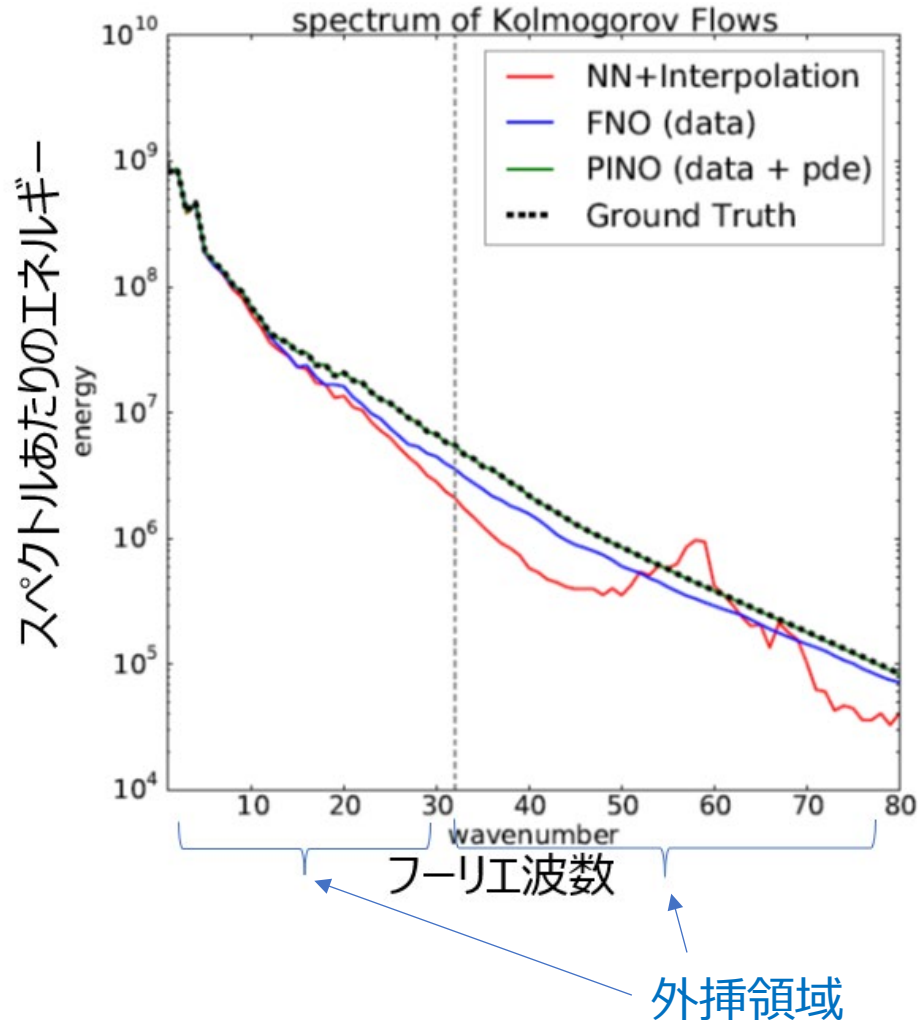
<https://arxiv.org/abs/2111.03794>, 2021

<https://dl.acm.org/doi/abs/10.1145/3648506>

ACM/JMS Journal of Data Science, 2024



# PINO vs. FNO



PINO(FNO+PINN)は、学習データと損失に対するPDEを制約に入れているため完全に復元できている。

# まとめ

データ駆動型アプローチ（機械学習）とモデル駆動型アプローチ  
を相補的に融合 = シミュレーションベース機械学習の紹介

深層学習は、これまでend-to-end学習で様々な応用分野で  
用いられてきたが、BPの自動微分機能を生かして、ODE,PDE  
の学習、求解などへの発展途上にある（AI研究の新潮流）