

A I による洪水予測と 水文観測データの異常検知

AIによる科学技術、経済、社会、
人類の進歩に取り残されないように、

一般財団法人 河川情報センター
情報基盤整備部 部長 田所 正
情報基盤整備部 副参事 岡田 治

1

1 Why AI? 3度目のAIが注目される理由



C. Bishop,
Edinburgh, Microsoft



G. Hinton
Google



A. Ng
Stanford



Y. LeCun
NY, Facebook



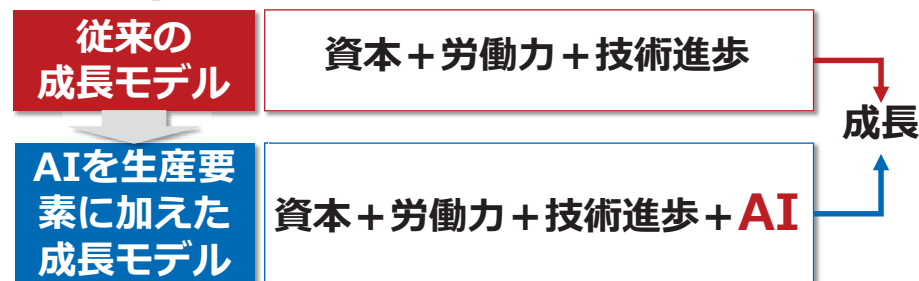
Y. Bengio
Montreal



I. Goodfellow
Google

ベイズ統計による機械学習が基礎

1 Why AI?

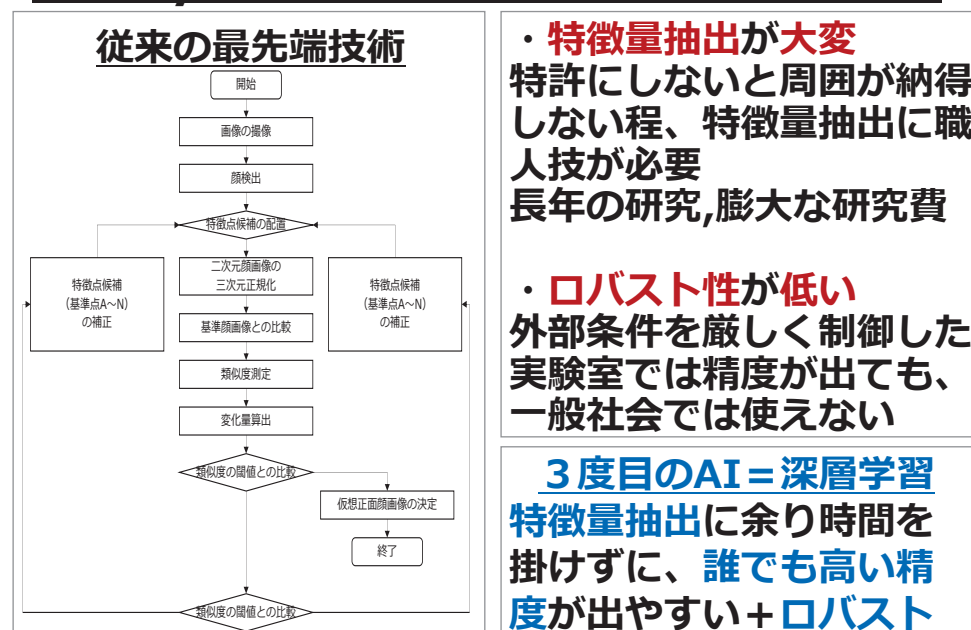


AIに積極的に取り組むことで、現在**0.8%**の成長率を**2035年**に**2.7%**の成長率へと押し上げると期待。現在、**600兆円/年**の経済規模に**+200兆円/年**を積み増す計算になる。**経済規模が倍**になるのに**80年**程を要していたが、**25年**程に短縮。

出典：Mark Purdy and Paul Daugherty, Why Artificial Intelligence is the Future of Growth, accenture, 2015

2

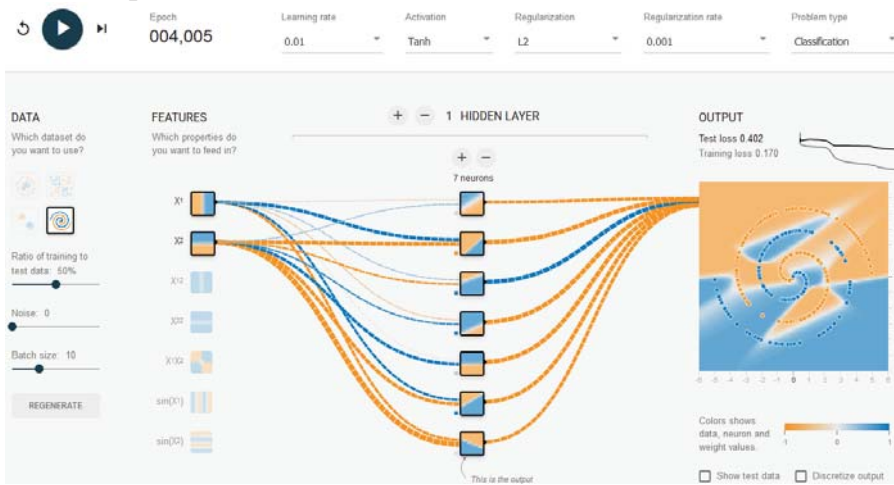
1 Why AI? 従来の技術と何が違う? ①



出典：顔認識装置及び顔器官の特徴点特定方法、KDDI、2009年登録済特許

4

1 Why AI? 従来の技術と何が違う? ②

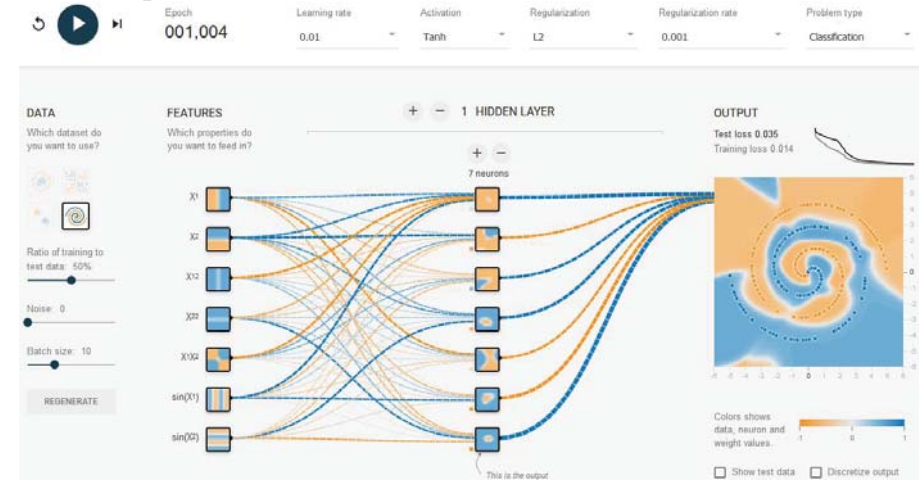


特徴量変換無しでシンプル, 浅層な AIでは分類 = 認識できない

出典: <https://playground.tensorflow.org/>

5

1 Why AI? 従来の技術と何が違う? ③

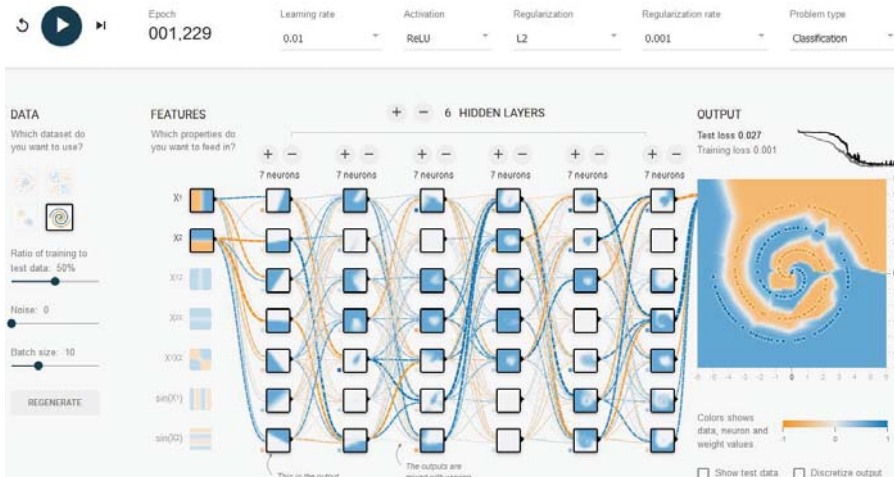


“適切”な特徴量変換 = 表現方法の工夫で分類可 (従来の認識技術と同様)

出典: <https://playground.tensorflow.org/>

6

1 Why AI? 従来の技術と何が違う? ④

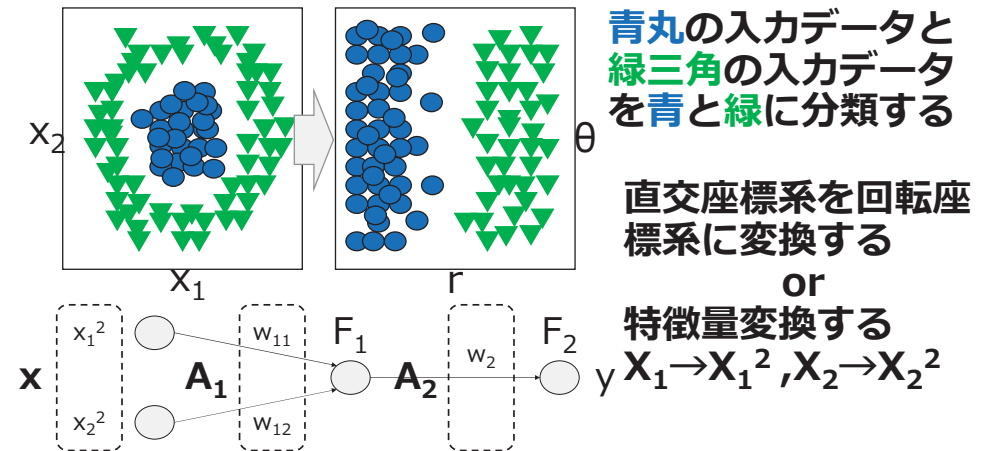


3度目のAI = 深層学習は“適切な”特徴量変換無しで動く, 研究室を出て実社会へ

出典: <https://playground.tensorflow.org/>

7

1 Why AI? AI内部で何が? ①



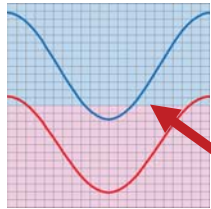
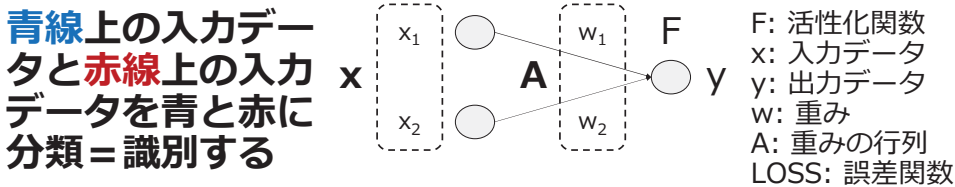
AI内部で行列 A_1 で平面の座標系を変換

出典: Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press, 2016, <http://www.deeplearningbook.org/>

8

1 Why AI? AI内部で何が? ②

青線上的入力データと赤線上的入力データを青と赤に分類 = 識別する



出力層で直線で分割

$$\operatorname{argmin}_A (\text{LOSS}(F(A, x), y))$$

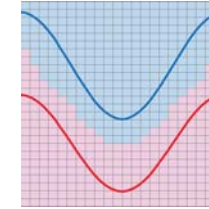
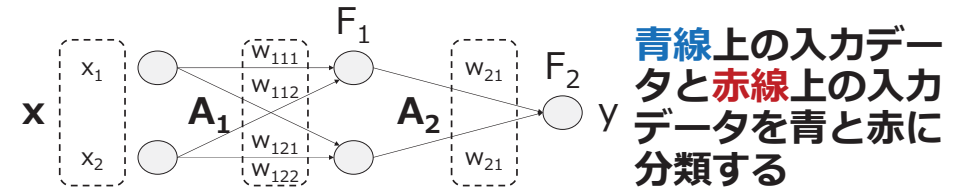
この式の意味: $\text{LOSS}(F(A, x), y)$ を最小化する行列 A

ニューロン 1 個の出力層でしていること
線分上で閾値で分類 = 認識

出典: Christopher Olah, Visualizing Representations: Deep Learning and Human Beings, 2015, colah.github.io/posts/2015-01-Visualizing-Representations/

9

1 Why AI? AI内部で何が? ③



入力

隠れ層

$$\operatorname{argmin}_{A_1, A_2} (\text{LOSS}(F_2(A_2, F_1(A_1, x)), y))$$

この式の意味: $\text{LOSS}(F_2(,,,))$ を最小化する行列 A_1, A_2 を求める

出力層で直線で分割

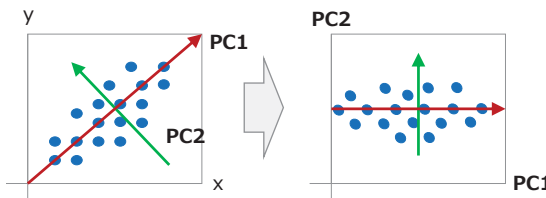
AI内部で行列 A_1 で平面の座標系を変換

出典: Christopher Olah, Visualizing Representations: Deep Learning and Human Beings, 2015, colah.github.io/posts/2015-01-Visualizing-Representations/

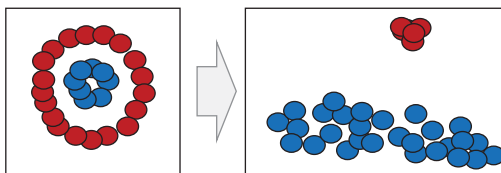
10

1 Why AI? AI内部で何が? ④

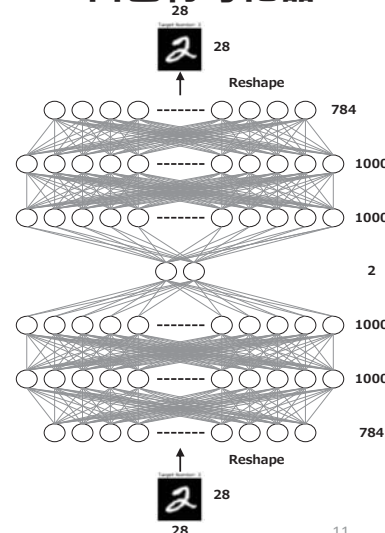
座標変換の従来技術
主成分分析: PCA



カーネル主成分分析: KPCA



3度目のAI = 深層学習
例: AutoEncoder
自己符号化器



出典: <http://scikit-learn.org>

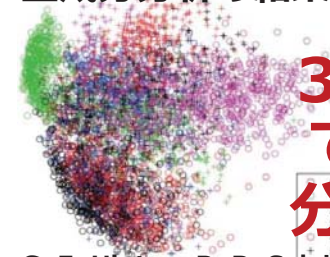
11

1 Why AI? AI内部で何が? ⑤

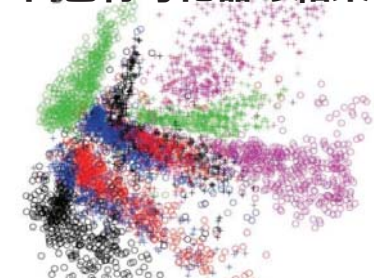
MNIST: 手書き数字
訓練55,000, テスト
10,000, 検証5,000

yann.lecun.com/exdb/mnist/

PCA
主成分分析の結果



3度目のAI = 深層学習
AutoEncoder
自己符号化器の結果



3度目のAI = 深層学習
では座標変換が良い,
分類 = 認識精度が高い

出典: G. E. Hinton, R. R. Salakhutdinov, Reducing the Dimensionality of Data with Neural Networks, SCIENCE VOL313, 2006

12

1 Why AI? 3度目のAI=深層学習なぜ 上手く分類=認識できるのか? Linの仮説



Henry
Lin,
Harvard



Max
Tegmark,
MIT

答えはNature of Universe, 数学より物理にある
百万画素のグレースケール画像は, 256^{1,000,000}
パターンあり、猫、犬等を簡単に分類=認識できる。
自然界の事象はHamiltonian, Bayesの定理で記述可
Hamiltonianには以下の特徴がある

1: **low polynomial order**, 2: **locality**,
3: **symmetry**, 4: **markov性**

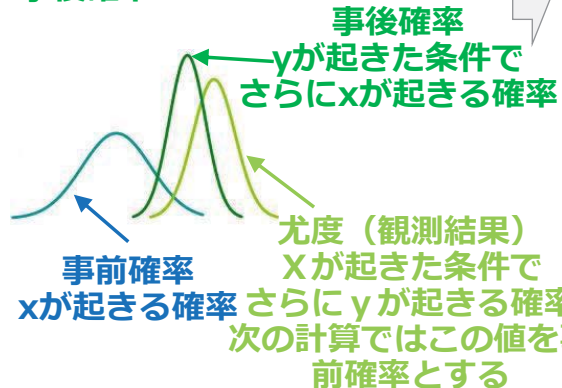
出典: H.Lin, M.Tegmark, Harvard, MIT, Why does deep and chep learning work so well?, 2016, arXiv:1608.08225v1

13

1 Why AI? 3度目のAI=深層学習なぜ 上手く分類=認識できるのか? Linの仮説

ベイズの定理 尤度 事前確率 ニューラルネットの各層
で出力される確率は
$$p(x|y) = \frac{p(y|x)p(x)}{\sum_{x'} p(y|x')p(x')}$$

事後確率 Hamiltonianで記述可



$$p(x|y) = \frac{1}{N(y)} e^{-[H(y)+u]}$$

Hamiltonianは多項
式で記述できる

$$H_x(y) = h + \sum h_{i1}y_i + \sum h_{ij}y_iy_j + \sum h_{ijk}y_iy_jy_k$$

出典: H.Lin, M.Tegmark, Harvard, MIT, Why does deep and chep learning work so well?, 2016, arXiv:1608.08225v1

15

1 Why AI? 3度目のAI=深層学習なぜ 上手く分類=認識できるのか? Linの仮説

Hamiltonian Mechanics: 例: 重力で落下する物体

$$H = KE + PE$$

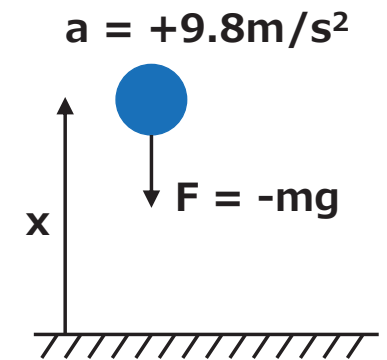
$$H = \frac{1}{2}mv^2 + mgx$$

$$= \frac{1}{2m} m^2 \dot{x}^2 + mgx$$

$$H = \frac{1}{2m} p^2 + mgx$$

$$\frac{\delta H}{\delta x} = mg \quad \frac{\delta H}{\delta p} = \frac{1}{2} p = v = \dot{x}$$

$$m \frac{d}{dt} \left(\frac{\delta H}{\delta p} \right) - \left(\frac{\delta H}{\delta x} \right) = 0$$



出典: H.Lin, M.Tegmark, Harvard, MIT, Why does deep and chep learning work so well?, 2016, arXiv:1608.08225v1

14

2 AIによる洪水予測 二瀬ダム流域



洪水調節を実施するには、**ダム流入量の予測が必要**
上流域が狭く、地上雨量計では**予測先行時間が短い**
降水量の予測には、WRF等モデルによる方法と**外挿**に
よる方法がある

本講演会では、**降水量予測AI (外挿)**、**降雨事例予測AI**、**ダム流入予測AI**を組み合わせる方法を説明する。

16

2 学習用データ

Cバンドレーダ雨量データ

降雨事例	洪水開始	洪水終了	洪水時間	ピーク雨量	総雨量
1	2006/10/04 22:00	2006/10/08 22:00	97	9.6	176
2	2007/09/04 23:00	2007/09/09 00:00	98	29.9	463.9
3	2011/09/02 00:00	2011/09/06 16:00	113	32.2	339.5
4	2011/09/19 16:00	2011/09/23 11:00	92	22.5	254.7
5	2012/06/19 03:00	2012/06/21 17:00	63	24.3	125.1
6	2013/09/15 17:00	2013/09/18 05:00	61	25.2	175.6
7	2014/10/04 22:00	2014/10/08 03:00	78	11.6	164.3
8	2015/07/15 17:00	2015/07/19 04:00	84	14.6	172.9
9	2015/09/07 21:00	2015/09/11 08:00	84	17.6	172.7

5分間隔、CSV形式データ

水文水質データベース

観測所番号	観測所名	水質項目	測定値
-------	------	------	-----

リアルタイムダム流量一覧表

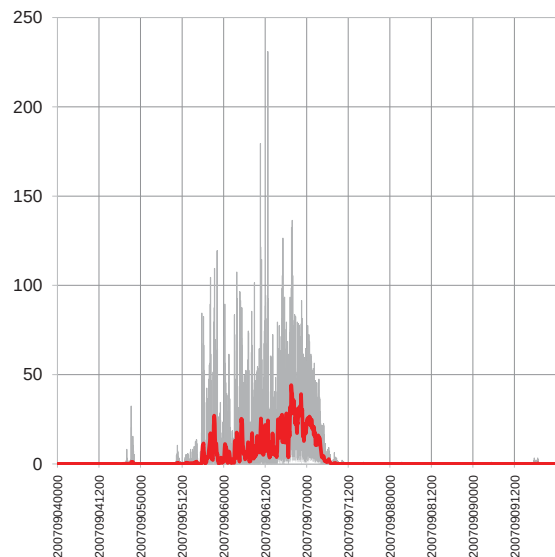
年月日	時刻	流量	水位	雨量	水温
2018/06/06	10:40	0.0	15400	1.38	4.01
2018/06/06	10:50	0.0	15400	1.38	4.02
2018/06/06	11:00	0.0	15400	1.38	4.04
2018/06/06	11:10	0.0	15407	1.42	4.02
2018/06/06	11:20	0.0	15407	1.42	4.01
2018/06/06	11:30	0.2	15407	1.42	4.03
2018/06/06	11:40	0.3	15407	1.42	4.02
2018/06/06	11:50	0.0	15407	1.42	4.02
2018/06/06	12:00	0.0	15414	1.47	4.00
2018/06/06	12:10	0.0	15414	1.47	4.04
2018/06/06	12:20	0.0	15414	1.47	4.04
2018/06/06	12:30	0.0	15414	1.47	4.01
2018/06/06	12:40	0.0	15414	1.47	4.01
2018/06/06	12:50	0.1	15422	1.47	4.00
2018/06/06	13:00	0.0	15422	1.47	4.02
2018/06/06	13:10	0.0	15422	1.47	4.00
2018/06/06	13:20	0.0	15422	1.47	4.02
2018/06/06	13:30	0.0	15426	1.47	4.04
2018/06/06	13:40	0.0	15426	1.47	3.98
2018/06/06	13:50	0.0	15426	1.47	4.00

10分間隔、CSV形式データ

水文水質データベースの流入量のデータは10分間隔。Pythonの関数でSpline補間し、5分間隔のデータに加工する。異常値、欠測値は、当該データを削除し、上記と同じくSpline補間する。

17

3 課題① レーダ雨量のノイズが大きい



降雨事例NO2：
2007年9月4日
縦軸：降水強度
横軸：1step=5分
灰線：Cバンドレーダ
同時刻合成雨量データ
赤線：雨量データの平均値

提案：
指数平滑化移動平均に
特徴量を変換する。

19

3 AIでダム流入量を予測するときの課題と提案

- ・レーダ雨量の**ノイズ**が**大きい**
- ・降雨からダム流入までの**時間**が**短い**
- ・降雨からダム流入までの**時間の幅**が**大きい**
- ・降水量と流入量の時系列変化の**形状**が**異なる**
- ・実運用に向け、**ロバスト**で**継続的**に**学習可能**

※書籍・チュートリアル・ブログ等に掲載の時系列予測のAIでは予測できません。

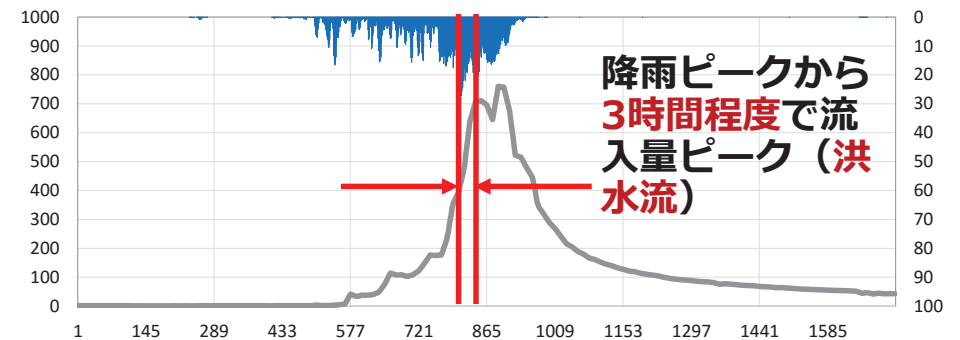
AI組み合わせ + 特徴量変換 + AI切替の手法を提案。

- ・0~6時間後の「**降水量を予測するAI**」
- ・類似の「**降雨事例を予測するAI**」
- ・「**ダム流入量を予測するAI**」

18

3 課題② 降雨からダム流入までの時間が短い

オリジナルのレーダ雨量データ等、降雨に早く反応する特徴量（ニューラルネットの入力値）が必要。

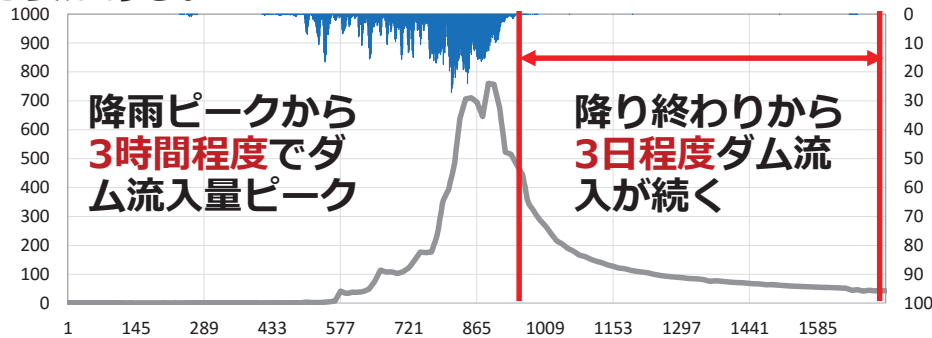


提案：指数平滑化移動平均に**特徴量**を変換
指数平滑化移動平均は移動平均より早く変化に反応
提案：レーダ雨量をAIで予測する。

20

3 課題③ 降雨から流入までの**時間の幅**が**大きい**

入力データのデータ数を3日以上とするか、降り始めからの累加雨量データをニューラルネットに入力する必要がある。

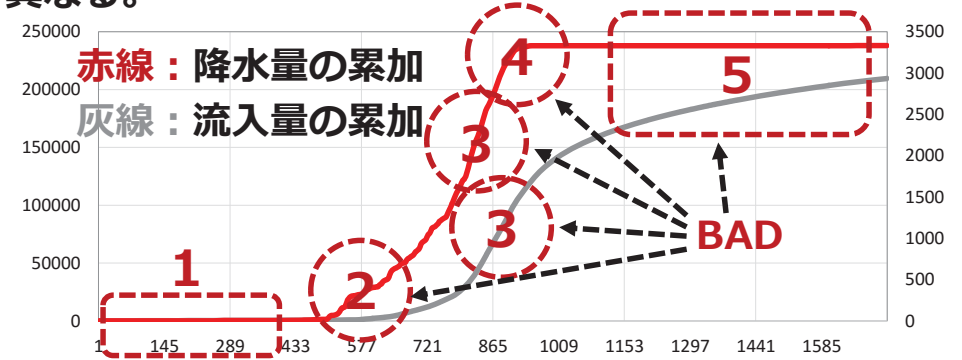


提案：3時間と3日間に同時に扱うため、**指数平滑化移動平均**を複数種類使う（240と2400タイムステップ）。

21

3 課題④ 降水と流入の時系列変化の**形状**が**異なる**

降水量と流入量を累加としても、時系列変化の形状が異なる。

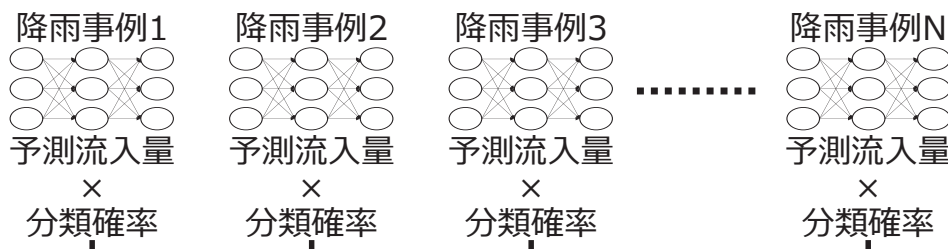


- ②降り始め、降水量の累加のみ上昇
- ③降水量と流入量の変曲点の時刻が異なる
- ④降り終わり、降水量の累加が大きく曲がる
- ⑤降水量の累加一定、流入量の累加、なだらかに上昇

22

3 課題⑤ 実運用に向け**堅牢**で**継続的**に**学習可能**

提案：時系列レーダ雨量の分類を**画像認識AI**で行う
提案：降雨事例毎に学習し、降雨事例毎の流入量の予測結果に分類の確率（0.0～1.0）を掛けて合成する。

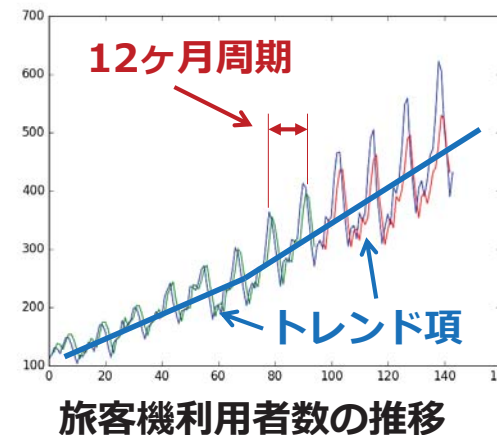


混合（アンサンブル）した予測流入量

降雨事例毎の学習では、**予測精度向上**が**容易**で、**洪水事例**の追加更新削除等、柔軟で**実運用**に近づく。

23

3 書籍等に掲載されているAI, **予測できません**

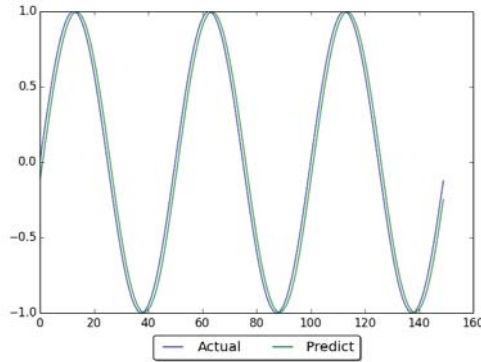


トレンド項、季節性（周期性）がある時系列データの予測の例。
ARIMA、SARIMAモデル等で内挿、外挿できる。下のURLの記事ではLSTMで1ステップ先を予測しているが、観測値の最後のデータを誤ってコピーし予測している。

<https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/>

24

3 書籍等に掲載されているAI, **予測できません**



LSTMによるSIN波の予測

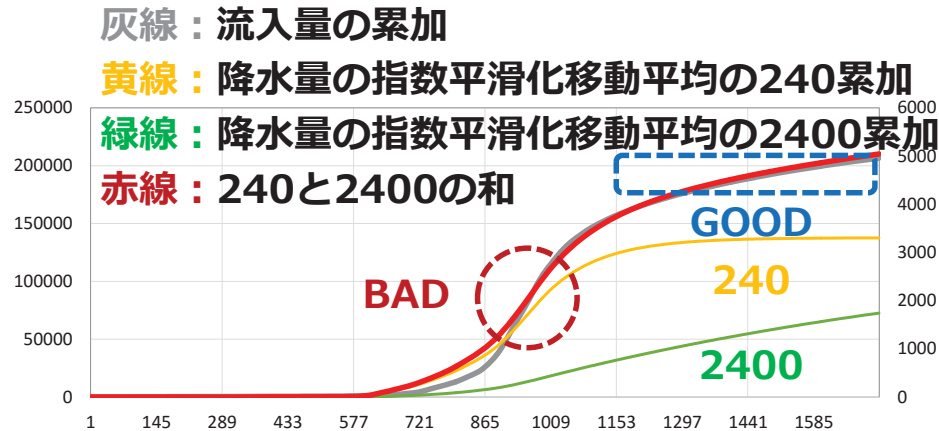
ブログ記事によると、Actual (実際) と Predict (予測) がずれているのは、線が重なるために意図的にずらした。

異なる波長のSIN波、 $y=x$ 等はこの学習済AIでは予測できない。現実の時系列データの予測 (株価、河川水位、流量の予測) では使えない。

<https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/>

25

4 本研究で提案する特徴量変換



降り終わり後は良く再現。

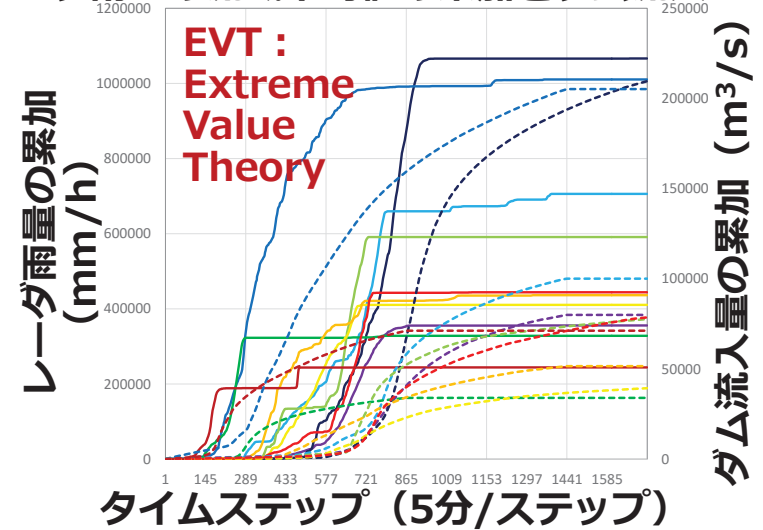
流入量ピーク付近が合わない (小さく予測)

提案 : 累加無しのニューラルネットも学習し、両者のMAX値を予測値として採用 (混合学習)

27

4 本研究で提案する特徴量変換

レーダ雨量の流域平均値の累加とダム流入量の累加



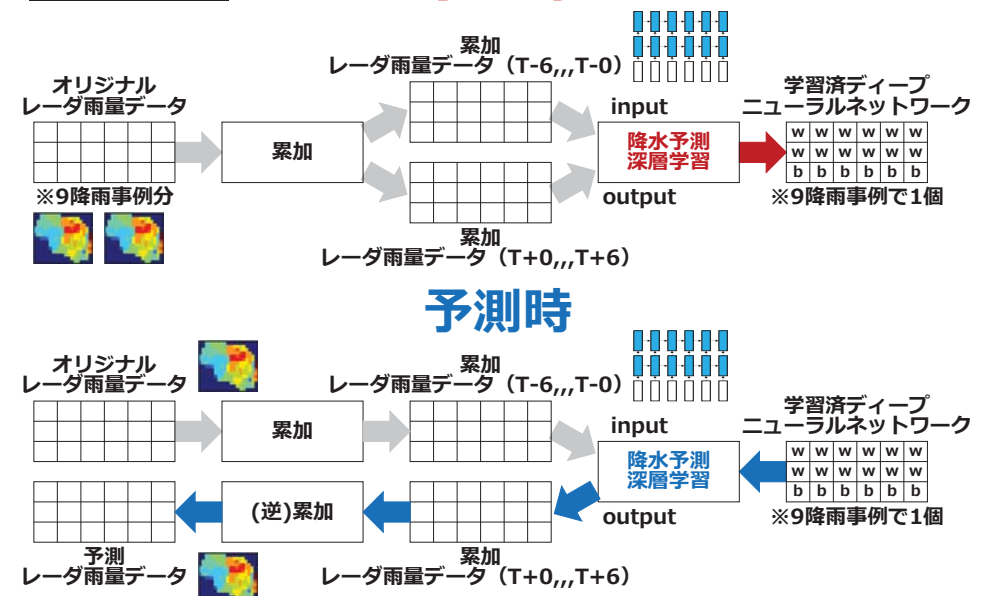
タイムステップ (5分/ステップ)

実線 : レーダ雨量の流域平均値の累加

破線 : ダム流入量の累加

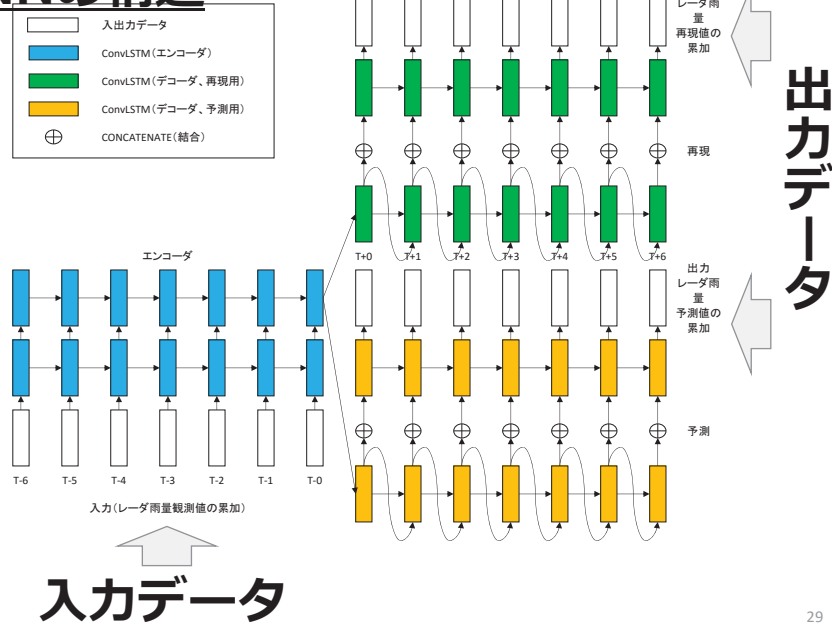
26

4 本研究で提案する降水量予測AIのアルゴリズム



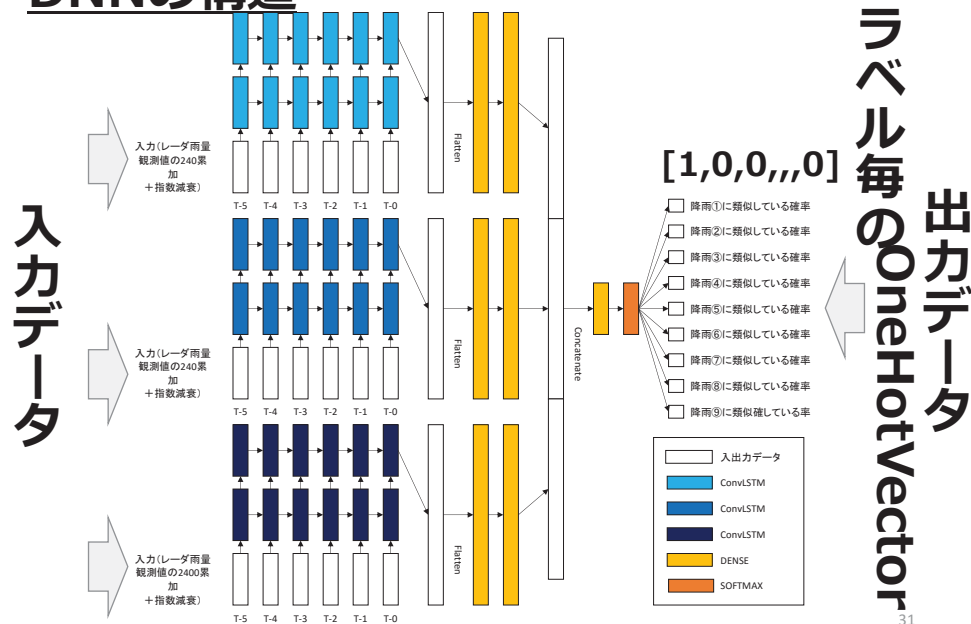
28

4 本研究で提案する降水量予測AIのDNNの構造



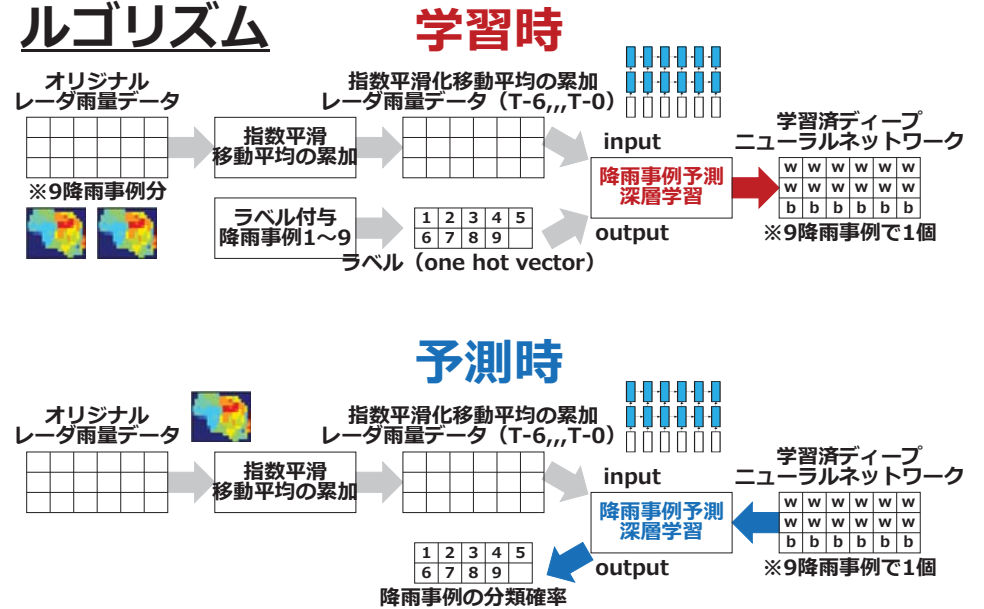
29

4 本研究で提案する降雨事例予測AIのDNNの構造



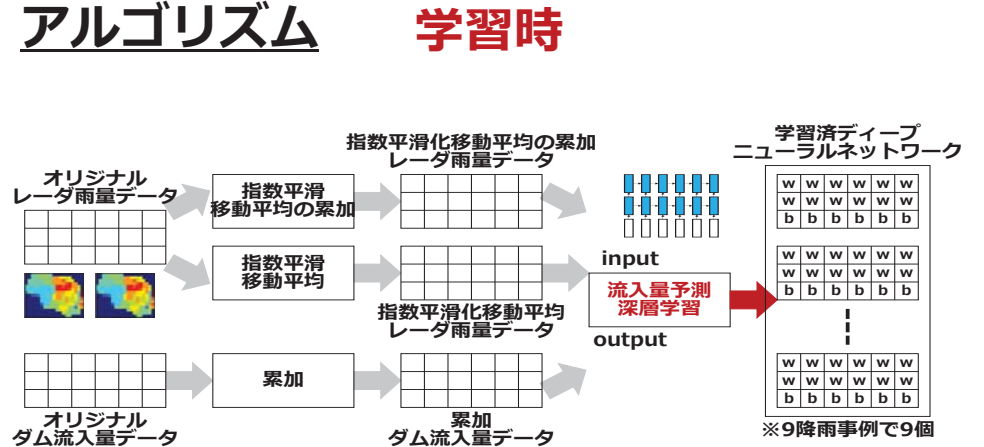
31

4 本研究で提案する降雨事例予測AIのアルゴリズム



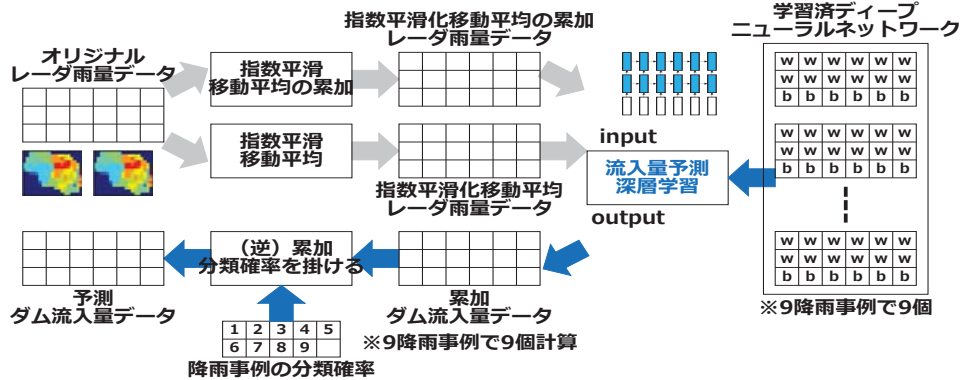
30

4 本研究で提案するダム流入量予測AIのアルゴリズム

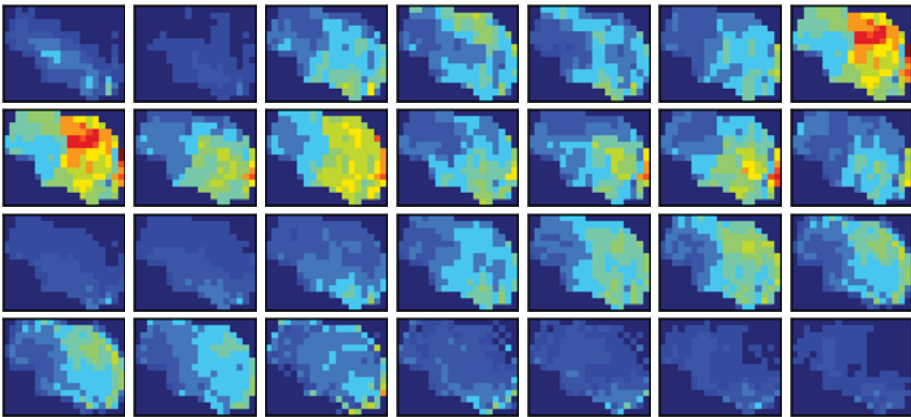


32

4 本研究で提案するダム流入量予測AIのアルゴリズム



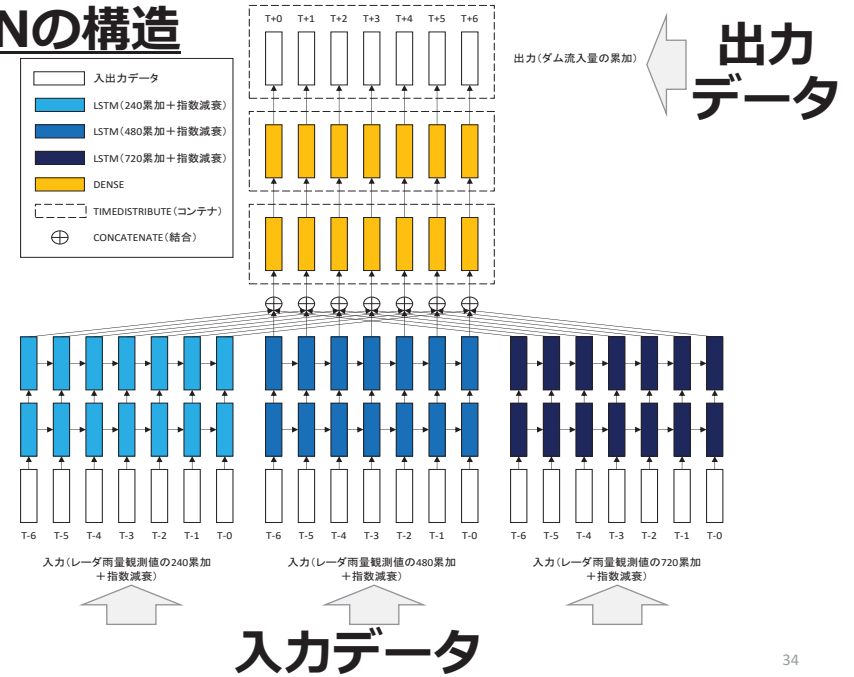
5 降水量予測AIの予測結果



降雨量予測AIの予測結果

上段：レーダ雨量の実測値、6、5、0時間前
2段：レーダ雨量の実測値、0、1、6時間後
3段：予測再現値、6、5、0時間前
下段：予測値、0、1、6時間後

4 本研究で提案するダム流入量予測AIのDNNの構造



5 降水量予測AIの精度

降雨事例（全9事例、下の図では例として5事例としている）を学習用と検証用に分け、**交差検証(Cross Validation)**により**予測精度**を評価する。

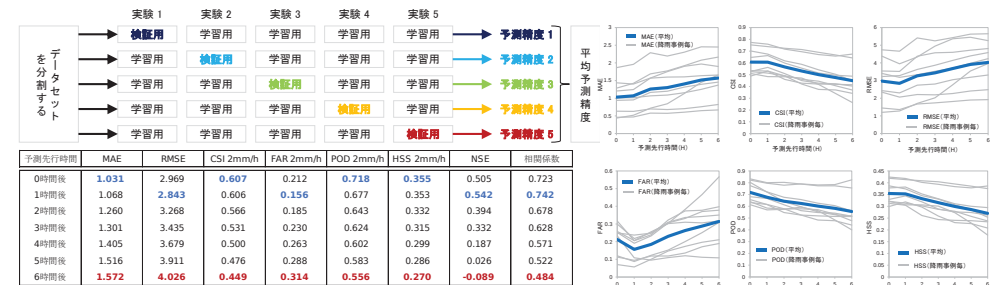


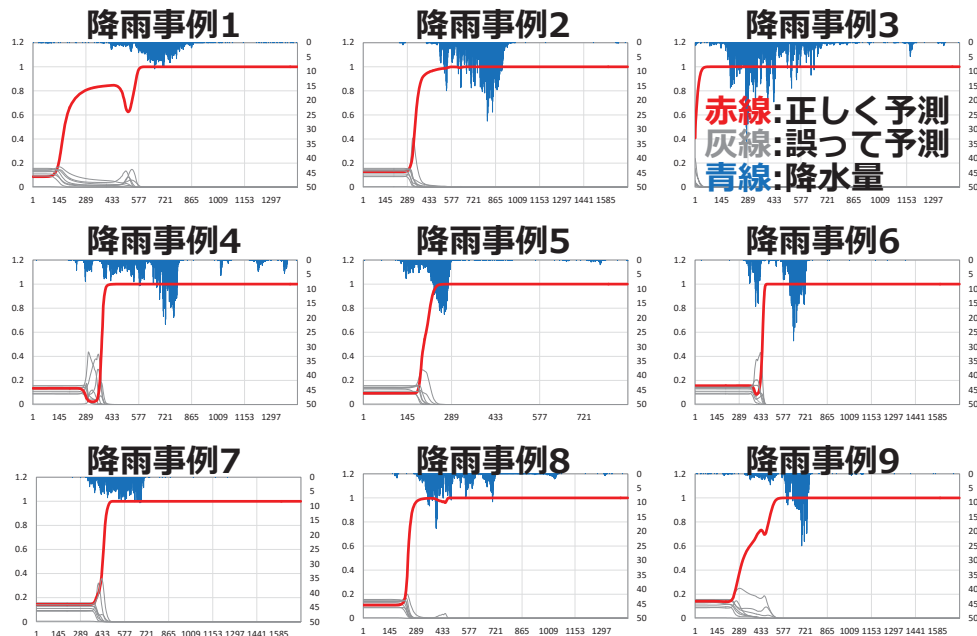
表 予測先行時間毎の予測精度

MAE: 小さいほど良い
RMSE: 小さいほど良い
CSI: 大きいほど良い
FAR: 小さいほど良い
POD: 大きいほど良い
HSS: 大きいほど良い
NSE: 大きいほど良い
相関係数: 1 に近いほど良い

予測先行時間(H) 予測先行時間(H)

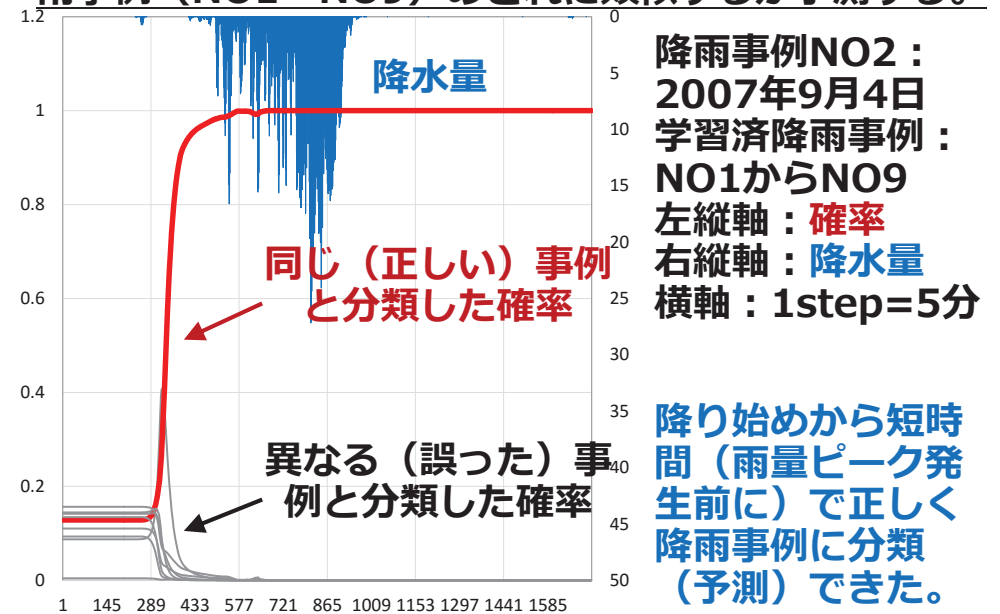
図 予測先行時間毎の予測精度

5 降雨事例予測AIの予測結果



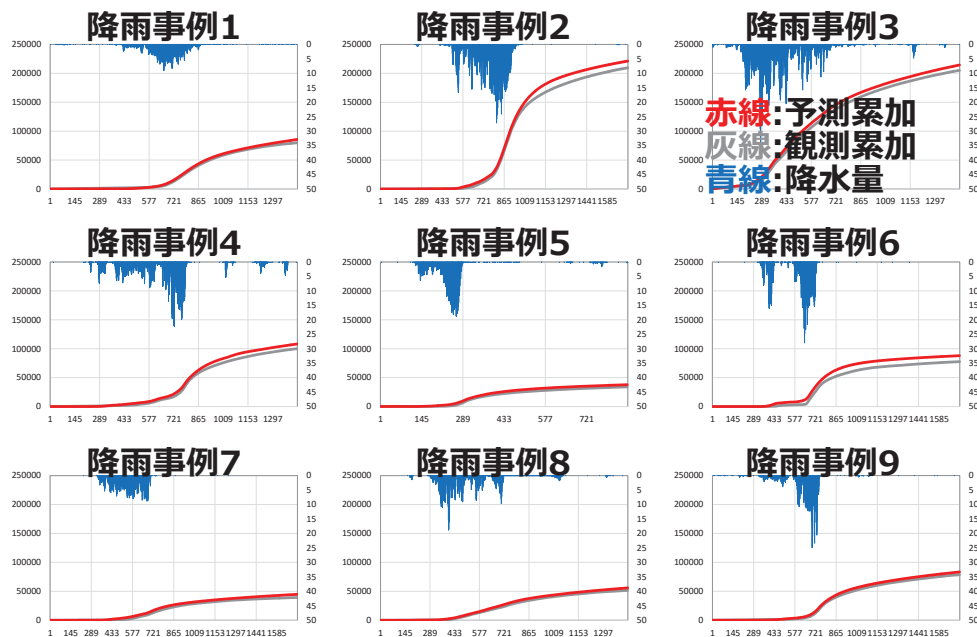
37

5 実験結果 予測対象の降雨事例NO2が学習済み降雨事例（NO1～NO9）のどれに類似するか予測する。



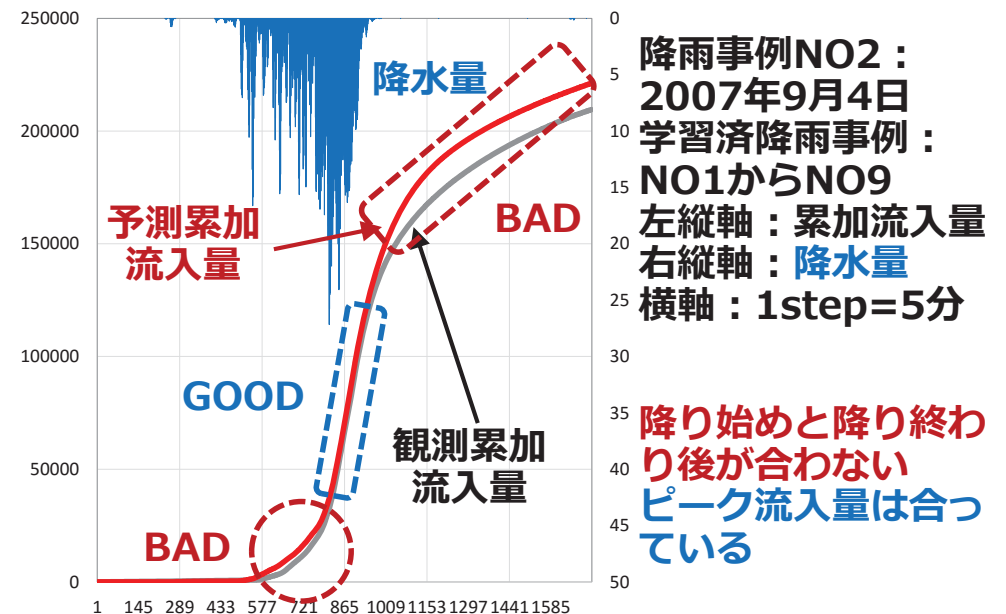
38

5 ダム流入量予測AIの予測結果（累加）



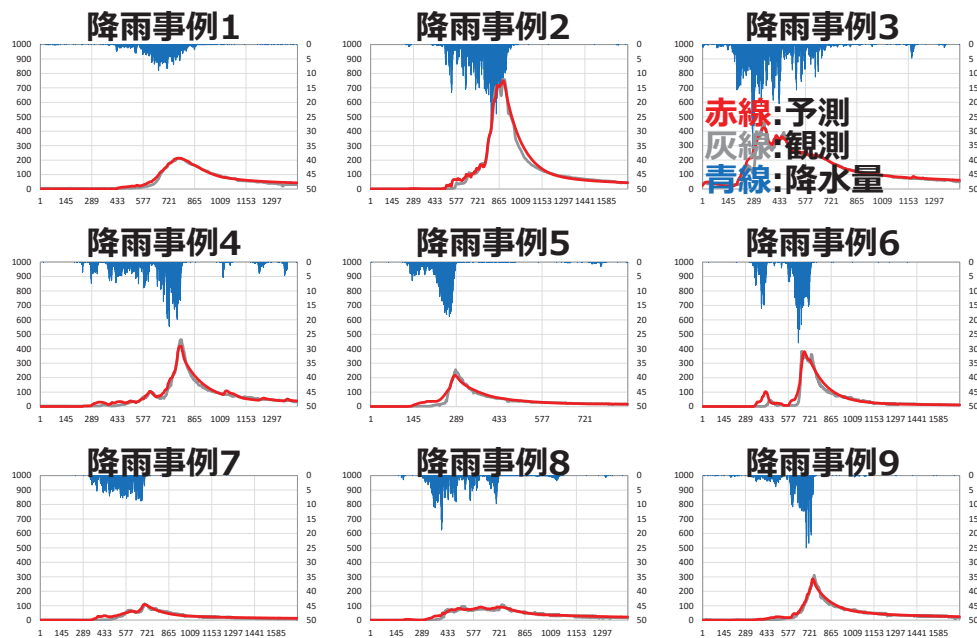
39

5 ダム流入量予測AIの予測結果（累加）



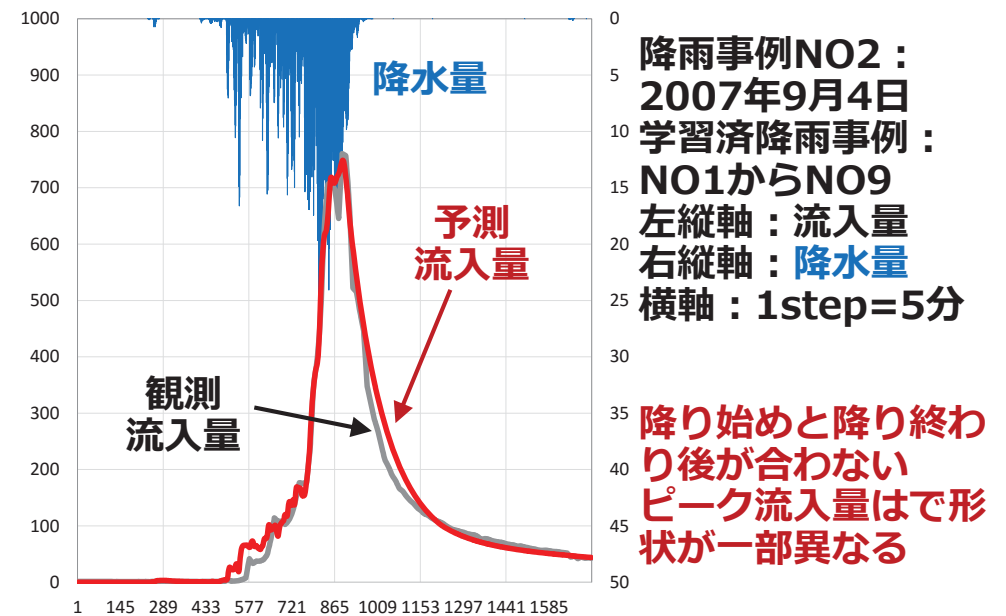
40

5 ダム流入量予測AIの予測結果



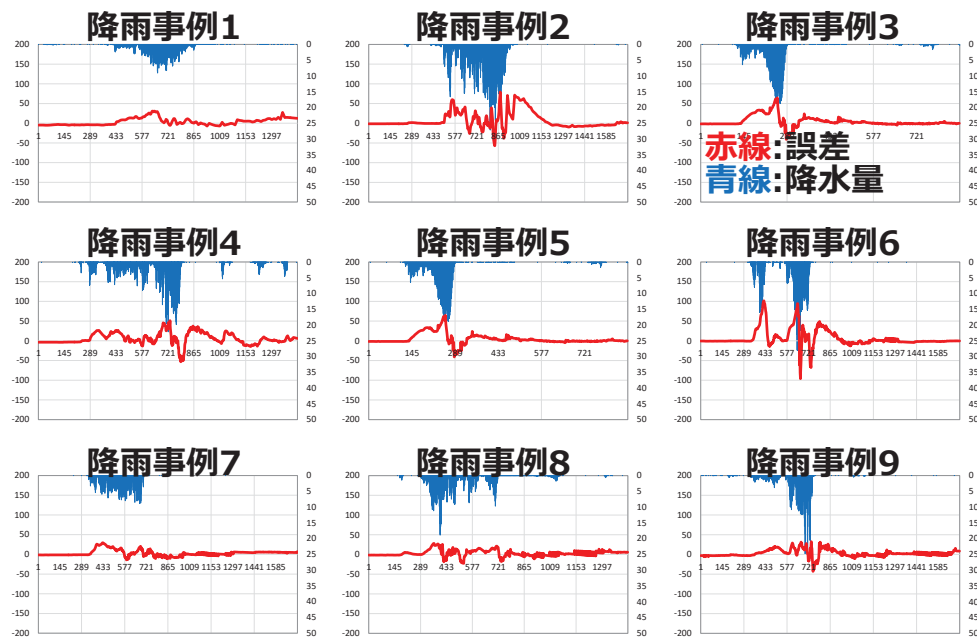
41

5 ダム流入量予測AIの予測結果



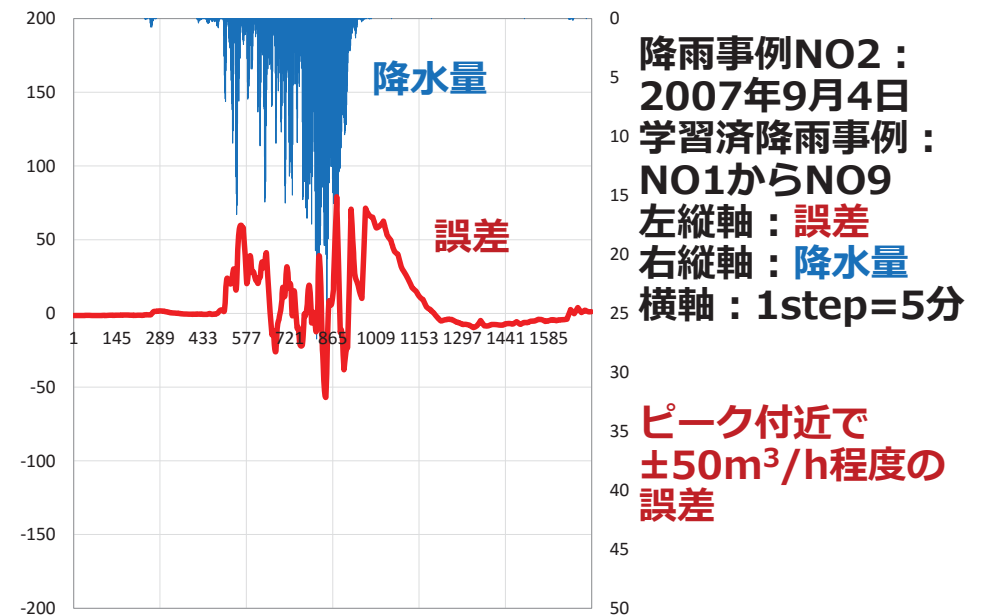
42

5 ダム流入量予測AIの予測結果（誤差）



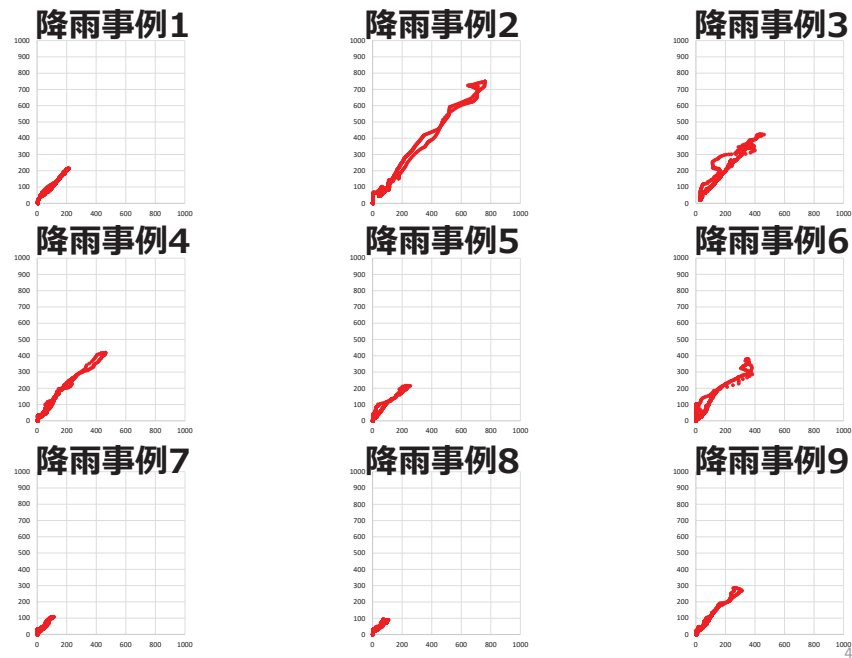
43

5 ダム流入量予測AIの予測結果（誤差）



44

5 ダム流入量予測AIの予測結果 (QQ)



45

5 実験結果のまとめ

降水量予測AI、降雨事例予測AI、ダム流入量予測AIを組み合わせ、従来より高い精度

今後、レーダ雨量の予測をより広い範囲で行い、ダム上流域の降水量の予測精度を高める

レーダ雨量の予測にD.Holden、S.Wiewelらのフェーズ、物理過程、雲の動きを組み込んだモデルを作成する

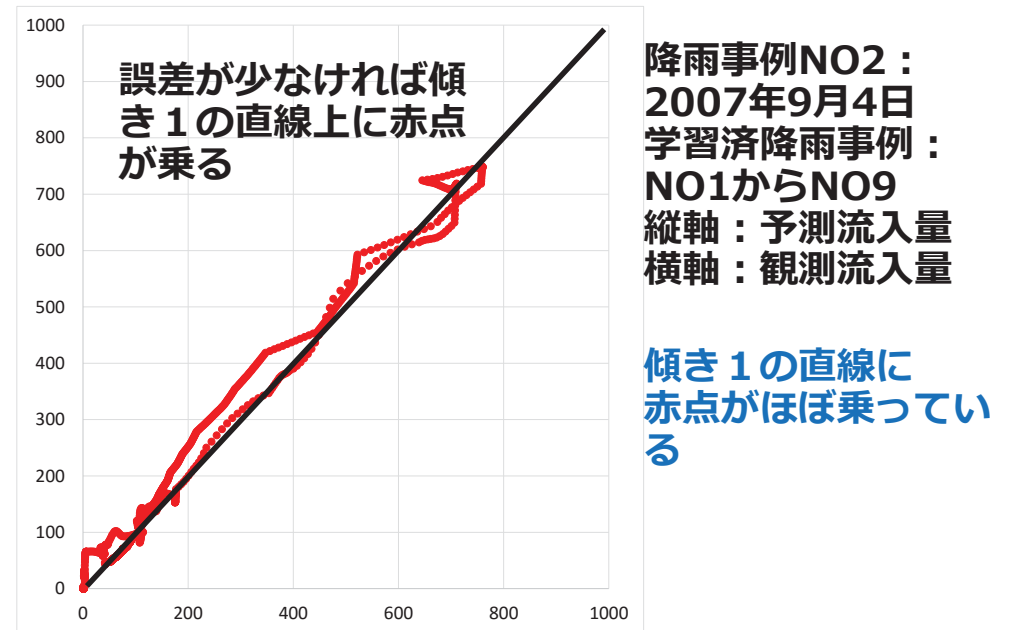
ダム流入量予測に降り始め時の土壌水分量に関するデータも考慮する

EVT : Extreme Value Theoryを使って100年、1000年に一度の降雨も想定内とする

WRF4.0の結果と組み合わせる

47

5 ダム流入量予測AIの実験結果 (QQ)



46

6 注目研究, AI関連技術

[AIの社会へのインパクト]

PwC, Stanford Woods Institute for the Environment, Harnessing Artificial Intelligence for the Earth, World Economic Forum, 2018
M. Purdy, Accenture, Why Artificial Intelligence is the Future of Growth

[動画像予測]

I.Sutskever, Sequence to Sequence Learning with Neural Networks, arXiv:1409.3215v3, 2014
N.Srivastava, Unsupervised Learning of Video Representations using LSTMs, arXiv:1502.04681v3, 2016
D.Holden, J.Saito, Phase-Functioned Neural Networks for Character Control, ACM Transactions on Graphics, Vol36, No.4, 2017
S.Wiewel, Latent-space Physics: Towards Learning the Temporal Evolution of Fluid Flow, SIGGRAPH2018, 2018
Deep Learning for Physical Processes: Incorporating Prior Scientific Knowledge, arXiv:1711.07970v2, 2018

48

6 注目研究, AI関連技術

[統計的ダウンスケール]

T.Vandal, DeepSD: Generating High Resolution Climate Change Projections through Single Image Super-Resolution, arXiv:1703.03126v1, 2017

T.Vandal, Quantifying Uncertainty in Discrete-Continuous and Skewed Data with Bayesian Deep Learning, arXiv:1802.04742v2, 2018

[データ同化]

K.Loh, Deep Learning and Data Assimilation for Real-Time Production Prediction in Natural Gas Wells, arXiv:1802.05141v2, 2018

大西領, 深層学習を活用した階層・分散型データ同化システム, 2017, 第2回理研データ同化ワークショップ

[レーダ雨量予測]

X.Shi, Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, arXiv:1506.04214v2, 2015

X.Shi, Deep Learning for Precipitation Nowcasting: A Benchmark and New Model, arXiv:1706.03458v2, 2017

49

6 注目研究, AI関連技術

[河川画像からの危険度抽出]

茨城大学とNEC, AI（人工知能）技術を活用した水害対策支援システムの実証実験を実施, 河川の画像から氾濫が発生する危険度の判断支援, NEC, 2017

[衛星画像からのセグメンテーション]

P.Helber, EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Cover Classification, arXiv:1709.00029v1, 2017

A.Albert, Using Convolutional Networks and Satellite Imagery to Identify Patterns in Urban Environments at a Large Scale, arXiv:1704.02965v2, 2017

51

6 注目研究, AI関連技術

[レーダ雨量画像ノイズ除去、補填]

S.Guo, Sea Clutter and Target Detection with Deep Neural Networks, 2nd AIEA, 2017

Q.Zhang, A Novel Method for Sea Clutter Suppression and Target Detection via Deep Convolutional Autoencoder, International Journal of Signal Processing, 2017

[レーダ画像からのExtreme Weather抽出]

Y.Liu, Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets, arXiv:1605.01156v1, 2016

E.Racah, Extreme Weather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events, 31st NIPS2017

E.Shi, Convolutional Neural Networks Applied on Weather Radar Echo Extrapolation, CSEA, 2017

A.McGovern, Using Artificial Intelligence to Improve Real-Time Decision-Making for High-Impact Weather, <https://doi.org/10.1175/BAMS-D-16-0123.1>

50

6 注目研究, AI関連技術

Boosting, Bagging

敵対的生成ネットワーク

カプセルネットワーク

ベイジアンニューラルネットワーク

最適輸送問題

WRF4.0

AWIPS2 CAVE, EDEX, IDV

6 書籍, YouTube

C.Bishop, パターン認識と機械学習, 岩波, 2007

I.Goodfellow, Y.Bengio, 深層学習, 2017

G.Hinton, NNs for Machine Learning, 2016

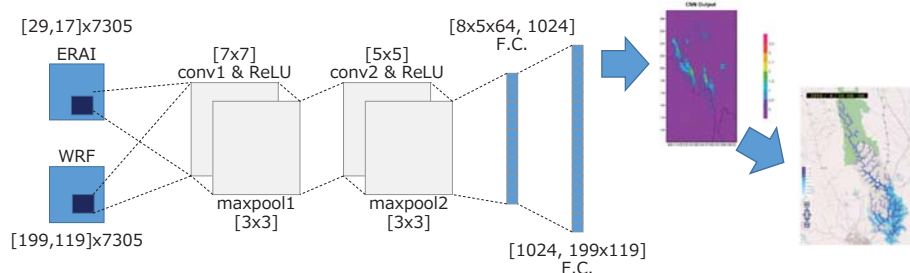
C.Manning, NLP with DL, 2017, Stanford Univ

L.Fridman, DL for Self-Driving Cars, 2017, MIT

52

6 注目研究, AI関連技術 統計的ダウンスケール

東大Y. Changらは全球気候モデル (Global Climate Model: GCM) を地域気候モデル (Regional Climate Model: RCM) に**ダウンスケール**するため、畳み込みニューラルネット (**CNN**) を使用する手法を提案。**RRI**で可視化。



衛星画像のセグメンテーションと洪水ハザードマップを組み合わせた。より深いDNNで精度向上をはかる。

出典 : Y. Chang, R. Acierto, A Deep Learning Approach to Downscaling Precipitation for Flood Risk Assessment over Data-scarce Resion 53

6 注目研究, AI関連技術 カプセルネットワーク

深層学習の父、現代の**アインシュタイン**(MIT TechReview)と言われる **G.Hinton**が**40年間**温め**2017年のNIPS**で発表
AIの中心と言える**CNNの問題点**を解消



Top: WGAN with MLP architecture. Bottom: Standard GAN, same architecture.

従来のCNN,
CNNを基礎としたGAN

出典 : S.Sabour, G.Hinton, Dynamic Routing Between Capsules, NIPS2017

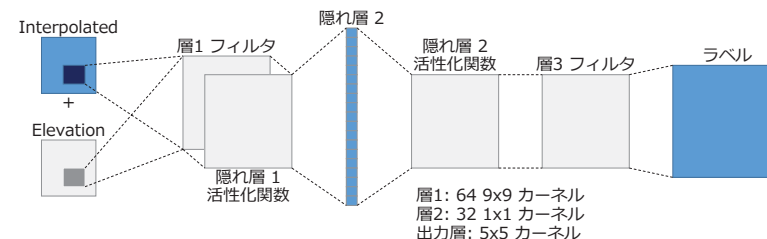
部分部分を学習するため、多層とすると画像が壊れる (目と耳の位置がピカソの絵)

1ピクセルの外れ値に影響され、キリンを自動車と誤認識する

プーリング層が悪い

6 注目研究, AI関連技術 統計的ダウンスケール

Northeastern Univ, T.Vandalらは、地球システムモデル (Earth System Models: ESMs) を地域気候モデル (Regional Climate Model: RCMs) に**ダウンスケール**するため、Chao Dongらが提案した Super-Resolution Convolutional Neural Network (**SRCNN**) を適用する手法を提案した。

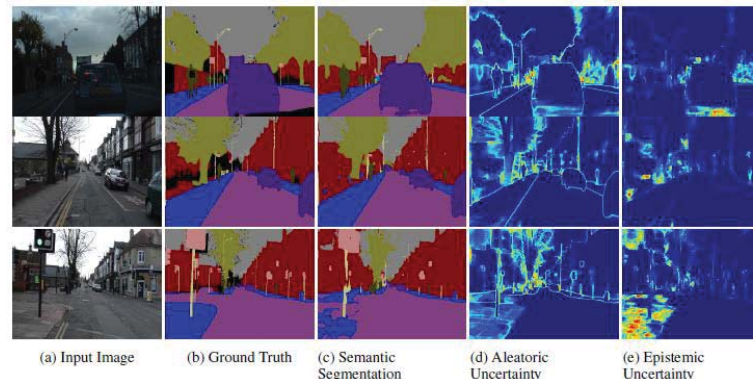


出典 : T. Vandal, DeepSD: Generating High Resolution Climate Change Projections through Single Image Super-Resolution, arXiv1703.03126
出典 : Chao Dong, 2014. Learning a deep convolutional network for image super-resolution., In European Conference on Computer Vision

6 注目研究, AI関連技術 ベイジアンDNN

Bayesian Deep Neural Network workshop at NIPS2017

確率分布を仮定したデータを入力すると、予測結果が**信頼範囲付き**で得られる。クリティカルな自動運転等、ノイズが大きい、数が少ないデータの学習にも有効。

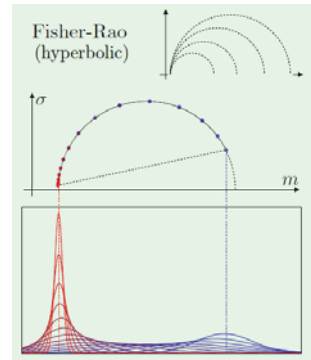


出典 : A.Kendall, What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?, arXiv:1703.04977v2, 2017
出典 : Y.Gal, Bayesian Deep Learning, Univ. of Cambridge, 2017

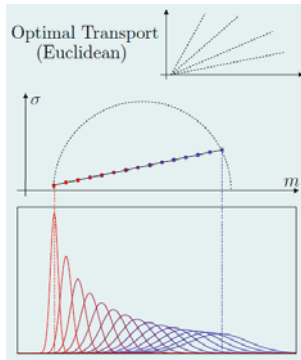
6 注目研究, AI関連技術 最適輸送

Optimal Transport Tutorial at NIPS2017

レーダ雨量画像が複数あった場合、その間を補間する際、**誤差関数でレーダ雨量画像の類似度を計算**する際に使えそう。



従来の手法



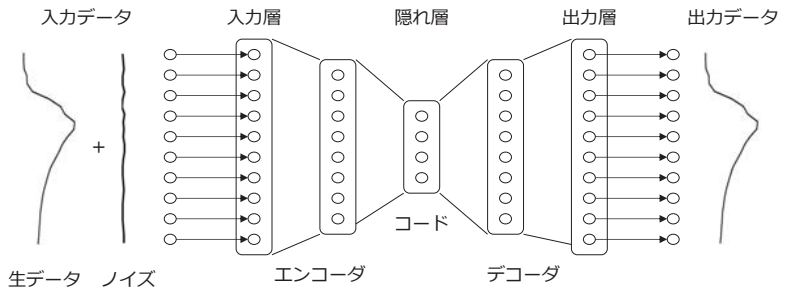
最適輸送

2つのガウス分布の間を補間する場合、従来の手法ではガウス分布自体は**移動せず、不自然な補間**になる。

元々は各地に散らばった兵員を目的地に輸送する最適輸送の計算に適用

出典：J.Solomon, Convolutional Wasserstein Distances: Efficient Optimal Transportation on Geometric Domains, NIPS workshop, 2017 57

7 AIを用いた水位データ異常検知のモデル



自己符号化器は、その構造に特徴がある。入力層が10ニューロンとすると、隠れ層4、出力層10等、隠れ層が絞られた構造を持つ。ラベル付き学習データの準備には多大なコストを要するが、入力層と出力層に同じデータ、例えば水位時系列データを入力し学習させるため、ラベル付けは必要ない。

ノイズ除去自己符号化器（Denoising AutoEncoder）は、AEの入力層にノイズを加えたデータを入力し、出力層で元のデータを再現させるように学習させる。

DAEの特徴は、**Low Dimension**な**High Degree**な表現を抽出できることにある。（参考、PCA：主成分分析）
この研究では、DAEを用い、水位データの異常を検出するモデルを提案する。

Ian Goodfellow, Yoshua Benjio, Deep Learning, MIT Press, 2016
<http://www.deeplearningbook.org> 58

7 AIを用いた水位データ異常検知の結果

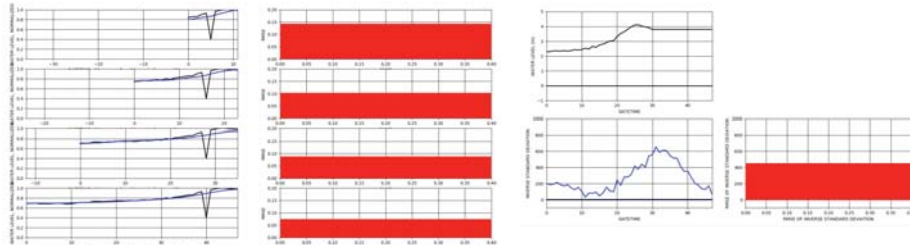


図 スパイクノイズの異常検知結果
左図：横軸時間、黒線：観測水位、青線：DAEの予測水位
右図：横軸時間、縦軸RMSE

図 長期無変動の異常検知結果
左図：横軸時間、黒線：観測水位、青線：DAEの予測水位
右図：横軸時間、縦軸RMSE

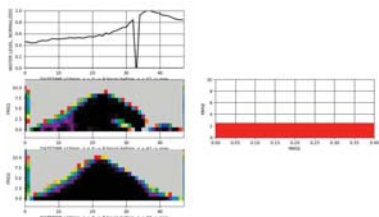


図 Wavelet変換したスパイクノイズの異常検知結果

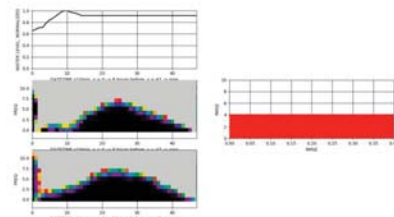


図 Wavelet変換した長期無変動の異常検知結果 59