

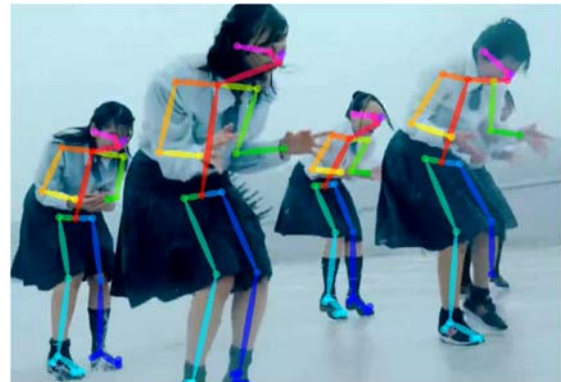
# ハイパパラメータ調整の自動化によるAI導入の加速

産業技術総合研究所 人工知能研究センター  
社会知能研究チーム  
大西 正輝

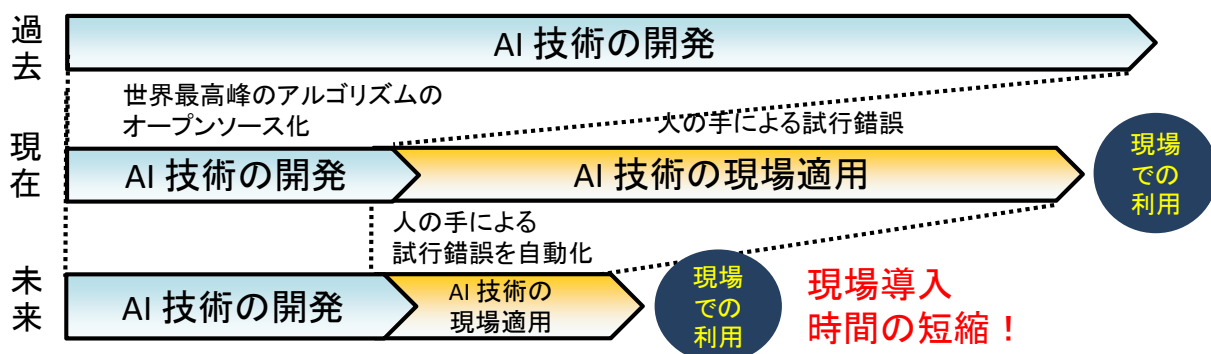
(連携希望) こんなことで困ってませんか？



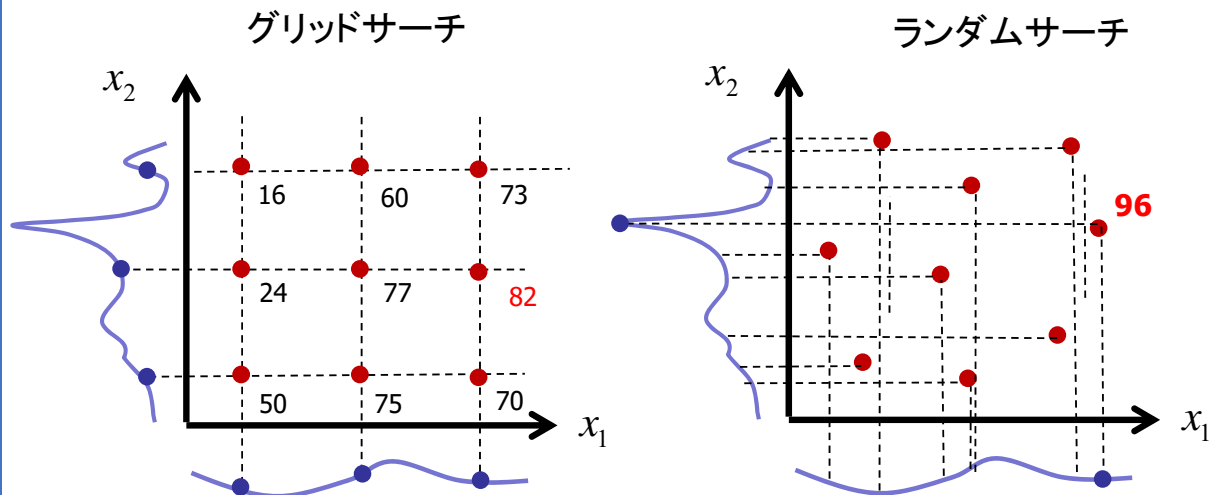
YOLO: Real-Time Object Detection



OpenPose: Real-time multi-person keypoint detection library for body, face, hands, and foot estimation



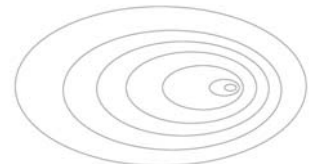
# ハイパラメータ調整とは？



それでも  
未だにグリッドサーチが使われている  
(機械学習の研究者でさえも・・・)

← ICML でゲットしたカップ

## 元になる論文発表



Osaki et al. *IPSJ Transactions on Computer Vision and Applications*, 2017, 9:20  
DOI 10.1186/s10174-017-0050-7

RESEARCH PAPER Open Access

Effective hyperparameter optimization using Nelder-Mead method in deep learning

Yoshihiko Ozaki<sup>1,2</sup>, Masaki Yano<sup>1,2</sup> and Masaki Onishi<sup>1,2\*</sup>

**Abstract**  
In deep learning, deep neural network (DNN) hyperparameters can severely affect network performance. Currently, such hyperparameters are frequently optimized by several methods, such as Bayesian optimization and the covariance matrix adaptation evolution strategy. However, it is difficult for non-experts to employ these methods. In this paper, we adapted the simpler coordinate-search and Nelder-Mead methods to optimize hyperparameters. Several hyperparameter optimization methods were compared by configuring DNNs for character recognition and age/gender classification. Numerical results demonstrated that the Nelder-Mead method outperforms the other methods and achieves state-of-the-art accuracy for age/gender classification.

**Keywords:** Hyperparameter optimization, Nelder-Mead method, Deep learning, Convolutional neural network

**1 Introduction**  
The evolution of deep neural networks (DNNs) has dramatically improved the accuracy of character recognition [1], object recognition [2, 3], and other tasks. However, their increasing complexity increases the number of hyperparameters, which makes tuning of hyperparameters an intractable task. Traditionally, DNN hyperparameters are adjusted using manual search, grid search, or random search [4]. However, search space expands exponentially relative to the number of hyperparameters; thus, such naive methods no longer work well. Therefore, more sophisticated hyperparameter optimization methods are required. In deep learning, a hyperparameter optimization problem can be formulated as a stochastic black box optimization problem to minimize a noisy black box objective function  $f(x)$ :  
Minimize  $f(x)$  ( $x \in \mathcal{D}$ ). (1)  
Here, using all information available about the objective function, we can obtain its value at point  $x$  with noise  $\epsilon$  as follows:  
 $y = f(x) + \epsilon$ ,  $\epsilon \sim \mathcal{N}(0, \sigma_{\epsilon}^2)$ . (2)

\*Correspondence: onishi@ipc.aist.go.jp  
<sup>1</sup>National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Tsukuba, Ibaraki, Japan  
<sup>2</sup>University of Tsukuba, 1-1-1 Tennoji, Tsukuba, Ibaraki, Japan

© The Author(s) 2017. **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Springer Open

[1] Yoshihiko Ozaki, Masaki Yano, Masaki Onishi, “Effective hyperparameter optimization using Nelder-Mead method in deep learning,” *IPSJ Transactions on Computer Vision and Applications*, (2017) 9:20.

CVA の Most accessed 4位 (2019年5月現在)

[2] 尾崎嘉彦, 矢野正基, 大西正輝, 久野誉人, “微分フリー最適化手法による識別器のハイパラメータチューニング,” 電子情報通信学会技術研究報告, PRMU2016-84, Sep. 2016.

2016年度PRMU研究奨励賞 受賞

[3] 渡邊修平, 尾崎嘉彦, 大西正輝, “Nelder-Mead法の並列化による識別器のハイパラメータチューニングの高速化,” 電子情報通信学会技術研究報告, PRMU2018-100, Jan. 2019.

2018年度PRMU研究奨励賞 受賞

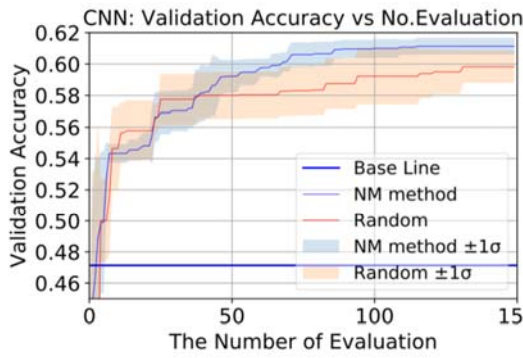
パラメータ調整の試行錯誤を自動化することで実用的な AI を速く作るための研究論文では深層学習を対象にしているが一般的な Black Box 関数の最適化に利用可

# 効果の検証①

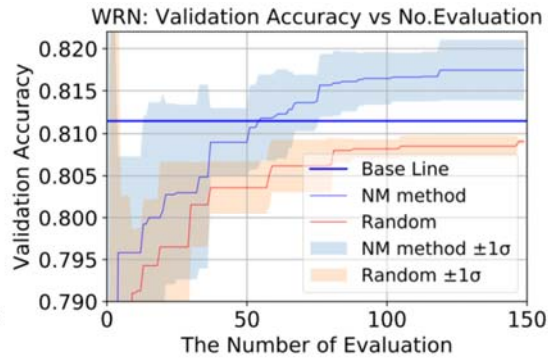
画像識別問題のデータセット (CIFAR-100) でNelder-Mead 法を検証



WideResNet (パラメータ数15) やDenseNetBC (パラメータ数101) で検証



Small CNN



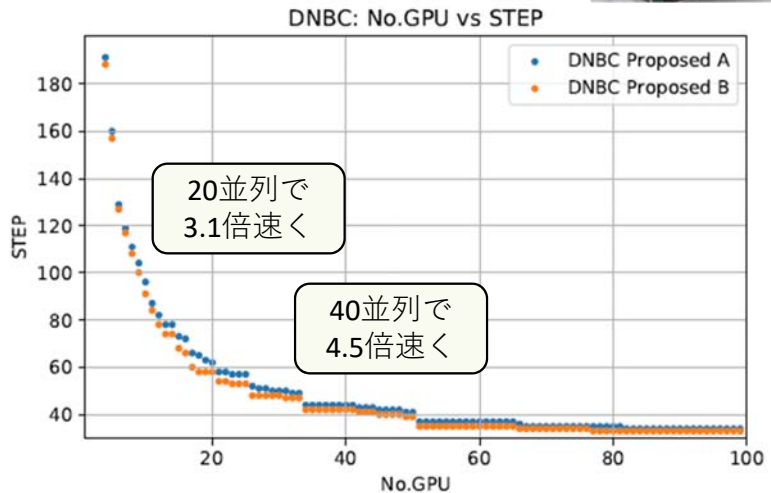
WRN

# 効果の検証②

**ABCI**  
All Bridging Cloud Infrastructure  
4352GPU, 43520CPU



ABCI を利用して投機実行による並列化の有効性を検証



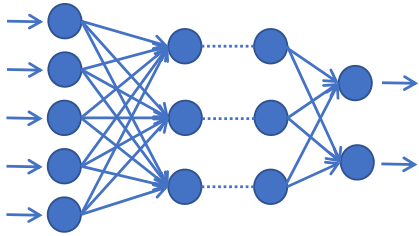
渡邊修平, 尾崎嘉彦, 大西正輝, “Nelder-Mead法の並列化による識別器のハイパーパラメータチューニングの高速化,”信学技報, PRMU2018-100, pp.33-38, Jan. 2019. PRMU研究奨励賞を受賞した手法をABCIに実装中 (2018年度に発表された170件の中から3件が受賞)



# 利用例 ①

## ◎ 高精度な深層学習の設計

Learning rate :  $\eta_0$    Momentum :  $\alpha$   
Weight decay :  $\lambda$    Batch size :  $B_t$   
.....



ハイパラメータ  $\mathbf{x} = \{\eta_0, \alpha, \lambda, B_t, \dots\}$   
が決まると

ネットワークの学習とテスト で

認識率  $f(\mathbf{x})$  が求まる

最も認識率の高い  $\mathbf{x}$  を知りたい

## ◎ ○○の良い××の設計

① ある程度次元数が高い問題が理想的 ( $n = 5 \sim 50$  程度)

②  $x$  は連続値を取ることが理想的

③ このプログラムはあることが前提

④  $f(\mathbf{x})$  の計算はある程度計算負荷の高い問題が理想的

ハイパラメータ  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$   
が決まると

計算機実験 で

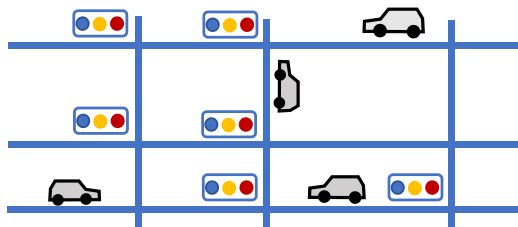
ある指標  $f(\mathbf{x})$  が求まる

最も指標の良い  $\mathbf{x}$  を知りたい

# 利用例 ②

## ◎ 混雑の少ない交通流制御の実現

信号制御 :  $S_n = \{G_n, Y_n, R_n\}$



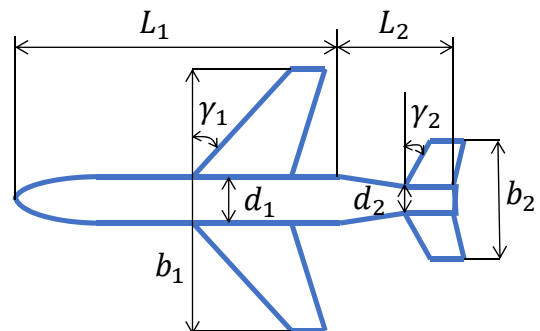
ハイパラメータ  $\mathbf{x} = \{S_1, \dots, S_n\}$   
が決まると

交通流シミュレーション で

混雑度  $f(\mathbf{x})$  が求まる

最も混雑度の低い  $\mathbf{x}$  を知りたい

## ◎ 燃費の良い飛行機的设计



ハイパラメータ  $\mathbf{x} = \{L_1, L_2, b_1, b_2, \dots\}$   
が決まると

流体力学のシミュレーション で

燃費  $f(\mathbf{x})$  が求まる

最も燃費の良い  $\mathbf{x}$  を知りたい



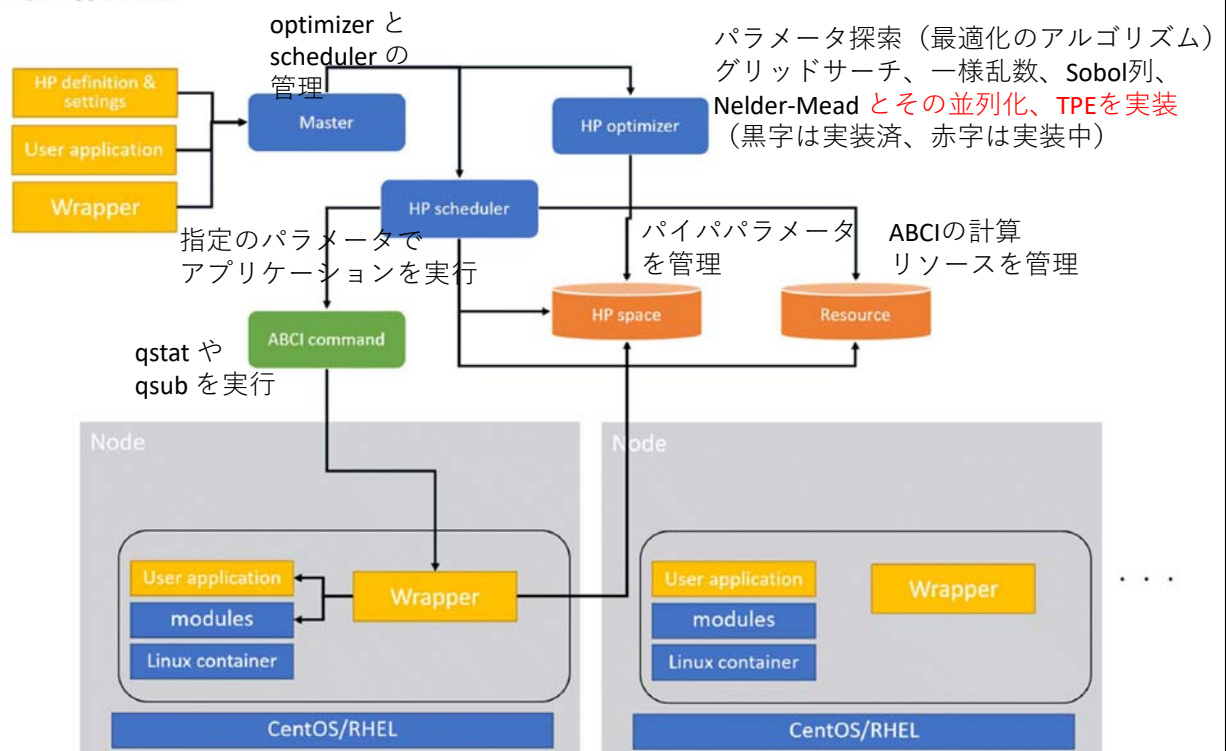
# OSSに向けて ①

- 産総研の AI クラウド (ABCI) を使って並列計算で速く精度の良い  $x$  を見つけるプログラム
- $x$  の値を決めると  $f(x)$  を求めるプログラムはあることが前提
- $f(x)$  を求めるプログラムのラッパーを作る必要がある
- ABCI だけではなくローカルマシンでも動作可能
  - ※ 並列計算ができないため最適化には時間がかかる
- ABCI では linux しか動作しないため windows でしか動作しない場合はローカルマシンで稼働
  - ※ ABCI では 4352GPU, 43520CPU を並列で利用可能

# OSSに向けて ②



で並列動作するようなフレームワークの基礎部分は完成



# 連携募集！

## ◎ ケース 1（ABCIを利用・作業は全て産総研）

- ・ プログラム改良費用と計算機使用料を産総研で負担
- ・ Linux で動作するソースコードを産総研に提供
- ・ 産総研の我々のチームでラッパーを作成
- ・ ABCIで最適化計算をしてパラメータセットを提供

## ◎ ケース 2（ABCIを利用・作業は全て御社）

- ・ 計算機使用料を産総研で負担
- ・ 産総研から御社に最適化プログラムを提供
- ・ 御社でラッパーを作成
- ・ ABCIで最適化計算をしてパラメータセットを取得

## ◎ ケース 3（作業は全て御社）

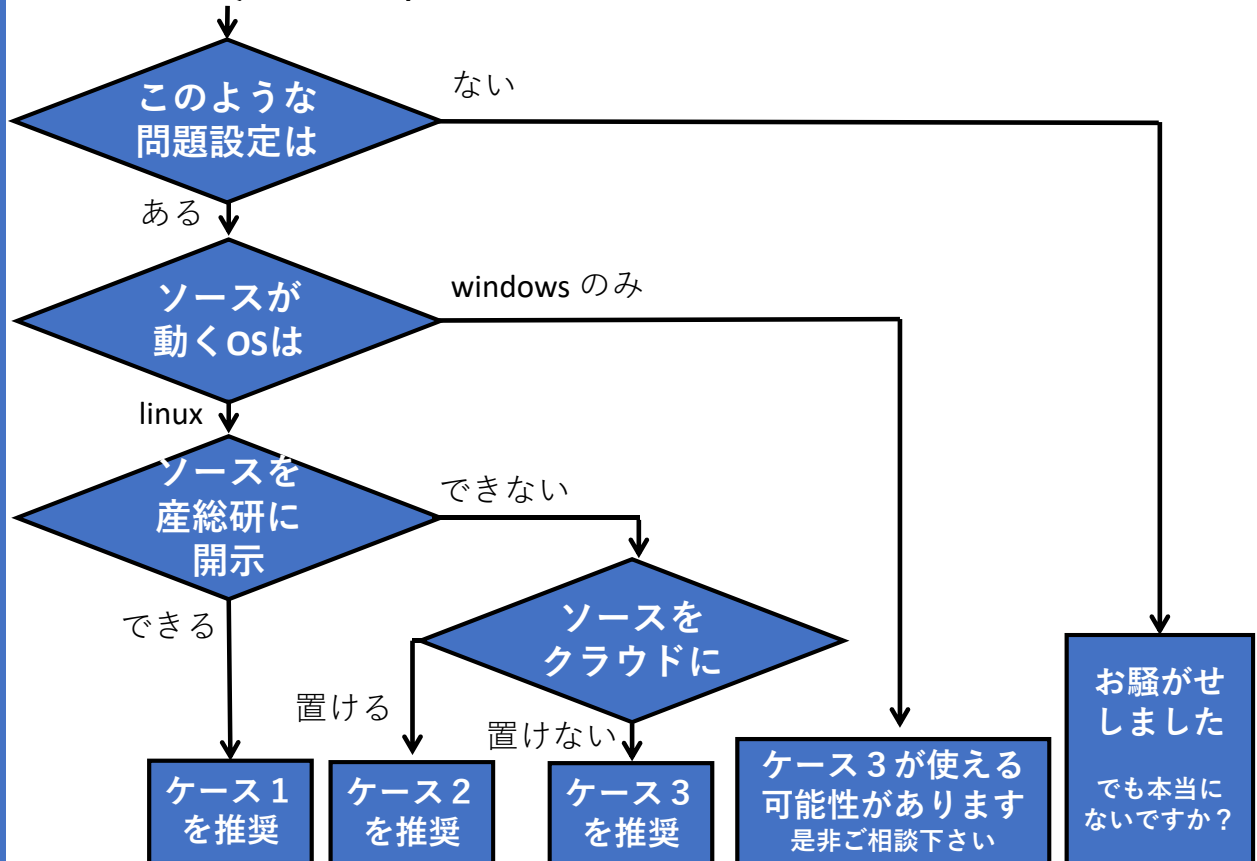
- ・ 産総研の負担はなし
- ・ 産総研から御社に最適化プログラムを提供
- ・ 御社でラッパーを作成
- ・ 御社のローカルマシンで最適化計算をしてパラメータセットを取得

※ 必要に応じて秘密保持契約

※ 必要に応じて御社社員を産総研で受け入れて作業可能

※ OSS 化を見越してこのソフトウェアを使いたいという連携も可能

# アンケート



# 連絡先

興味や質問がありましたら

[onishi@ni.aist.go.jp](mailto:onishi@ni.aist.go.jp)

にご連絡下さい  
もちろんこの後でも構いません