

企画セッション 「深層学習の理論」

2019/11/22 IBIS2019

企画者：今泉允聡

(統計数理研究所 / 理研 / JSTさきがけ)

導入：深層学習の登場

基礎研究

~2000

ブレイクスルー

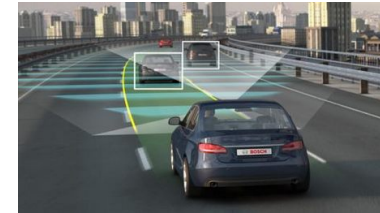
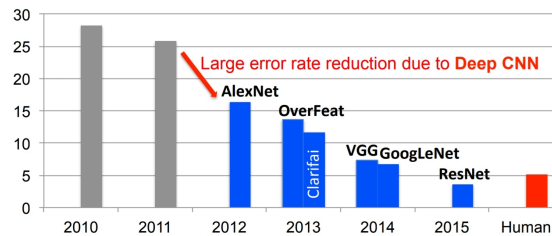
2012

実用化の進展

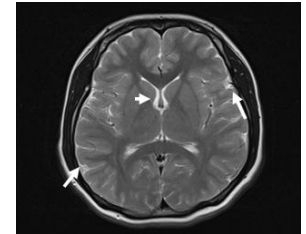
2016~

技術的
課題

技術
発展



自動運転^[1]



医療診断^[2]

高精度を発揮

画像識別精度: 75% → 96%

なぜDNNは高性能なのか？
⇒ **原理の解明はまだ発展途上**

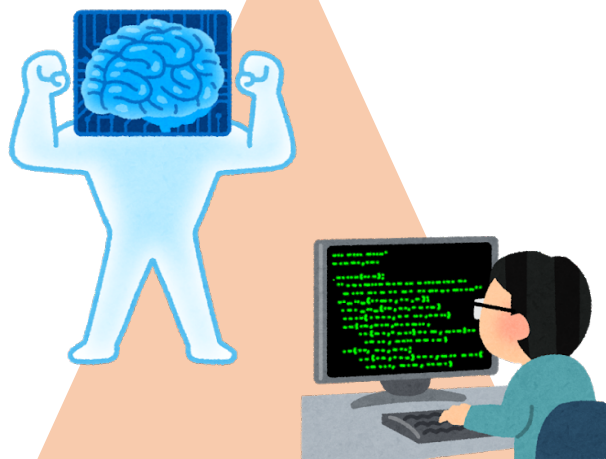
内部の解釈や効率的運用に向けて

発展する深層学習

実運用

- 発展するAI技術
- 賢いアルゴリズム

実運用



応用

基盤となるべき理論

- 実現象との橋渡しはある
- 今後の拡張が必要

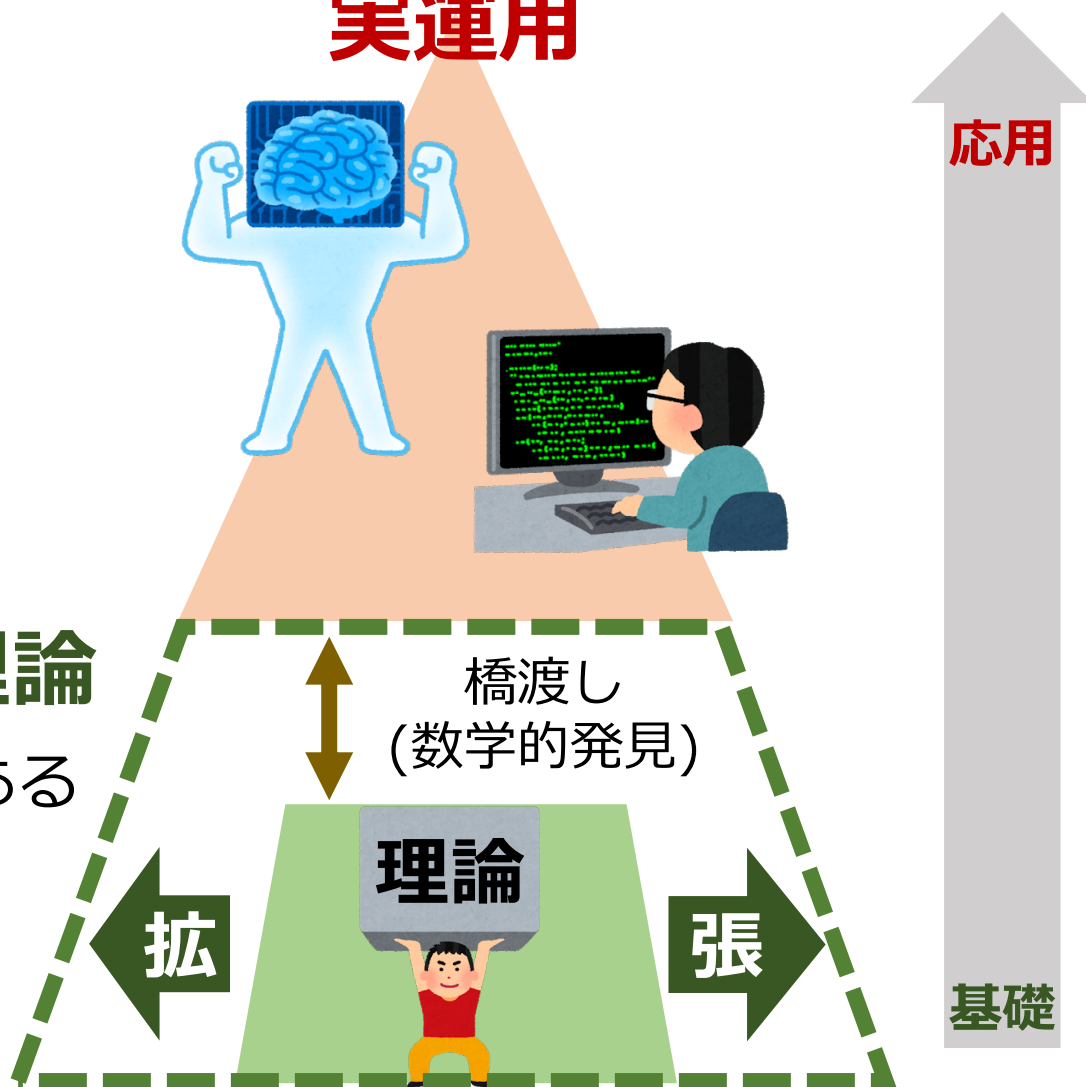
橋渡し
(数学的発見)

理論

拡

張

基礎



企画の概要

- 深層学習の精度(**汎化誤差**)を2側面から分析

1. 汎化誤差の構成要素

2. 特定構造



今日の
講演

近似/複雑性
最適化
群対称性

: 今泉允聡 (統数研)

: 二反田篤史 (東大)

: 三内顕義 (理研) (敬称略)

統一記法

- 設定
 - パラメータ Θ
 - DNNによる関数 $f_{\Theta}: \mathcal{X} \rightarrow \mathbb{R}$
 - 損失関数 $\ell(y, f_{\Theta}(x))$
 - 観測データ $D_n = \{(x_i, y_i)\}_{i=1}^n$
 - 経験損失 $\mathcal{L}(\Theta) = \frac{1}{n} \sum_{i=1}^n \ell(y_i, f_{\Theta}(x_i))$
- ニューラルネットワーク
 - レイヤー数 L
 - 活性化関数 σ
 - ℓ 層のパラメータ Θ_{ℓ}
 - ℓ 層の変換 $h_{\ell} = \sigma(\Theta_{\ell}^{\top} h_{\ell-1})$
- 勾配法
 - Θ による微分 $\nabla_{\Theta} \mathcal{L}(\Theta)$
 - SGDのステップサイズ η
 - SGDによるパラメータの更新 $\Theta^{(t+1)} = \Theta^{(t)} - \eta \nabla_{\Theta} \mathcal{L}(\Theta^{(t)})$