

次世代脳型人工知能技術のための 深層学習理論解析

奈良先端科学技術大学院大学

数理情報学研究室

池田 和司



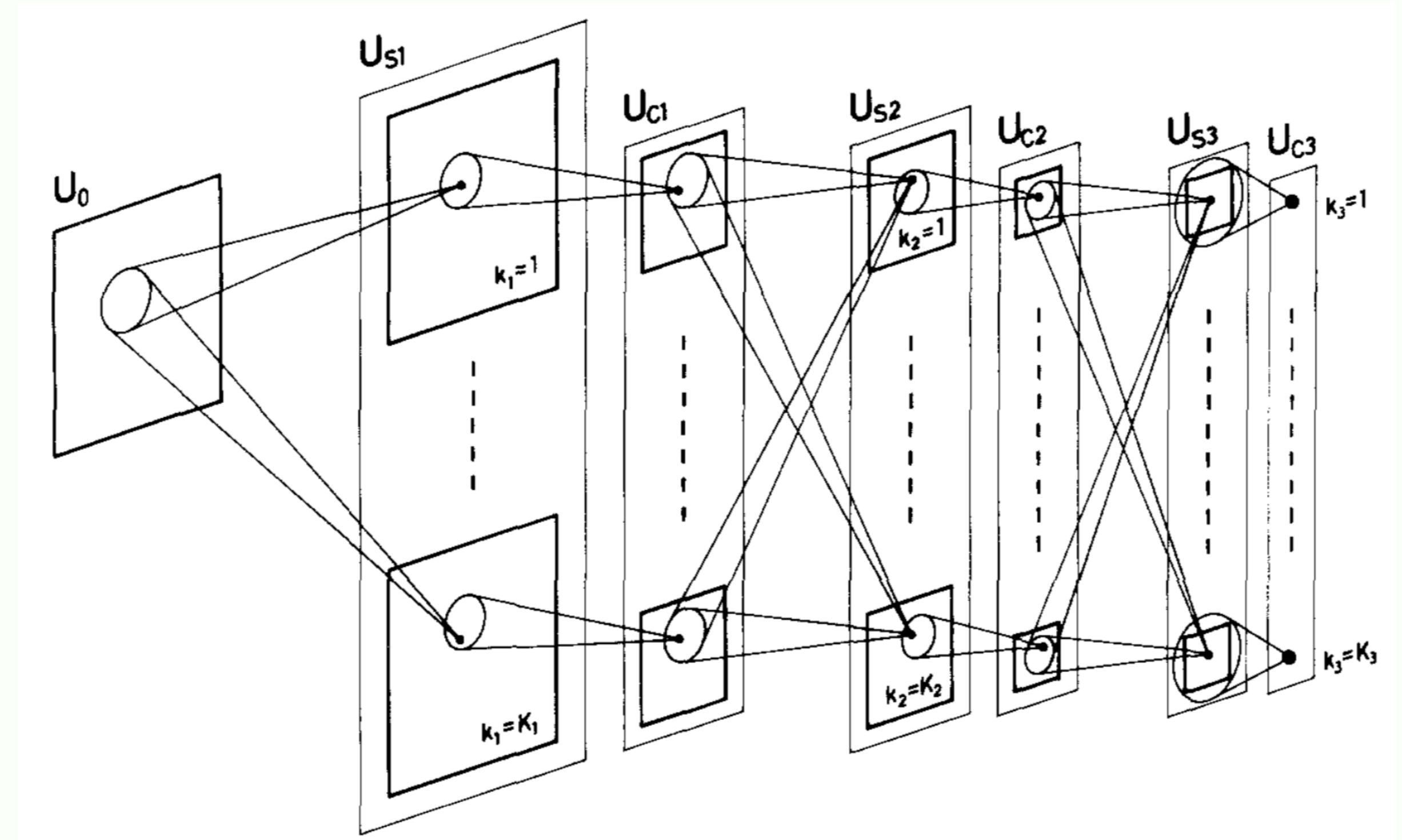
理論研究がニューラルネットの鍵

第1次ニューロブーム

- ・パーセプトロンの収束定理
- ・パーセプトロンの有界定理

第2次ニューロブーム

- ・バックプロパゲーション学習
- ・万能関数近似



(Fukushima, 1980)

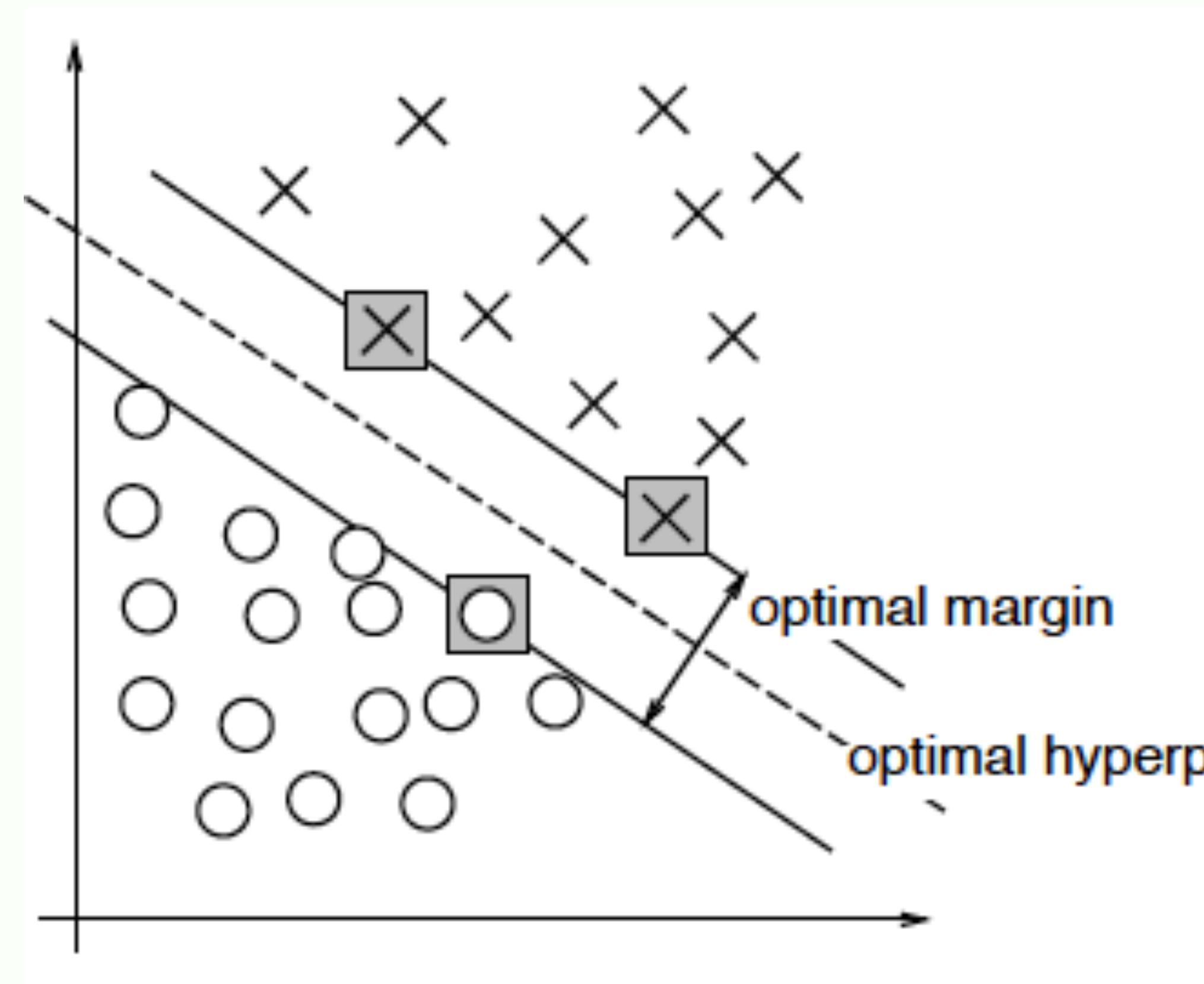
ブームの終焉も理論による

第1次ニューロブーム

- 線形分離可能性の要求
- ランダムな前処理は汎化力ゼロ

第2次ニューロブーム

- ローカルミニマとプラトー
- cf. SVMは唯一解
- 解の評価は計算機実験
- cf. SVMは学習理論による誤差の上限



(Cortes & Vapnik, 1995)

深層学習を超えるために深層学習を知る

確率降下法 (Robbins & Monro 1951)

プレトレーニング

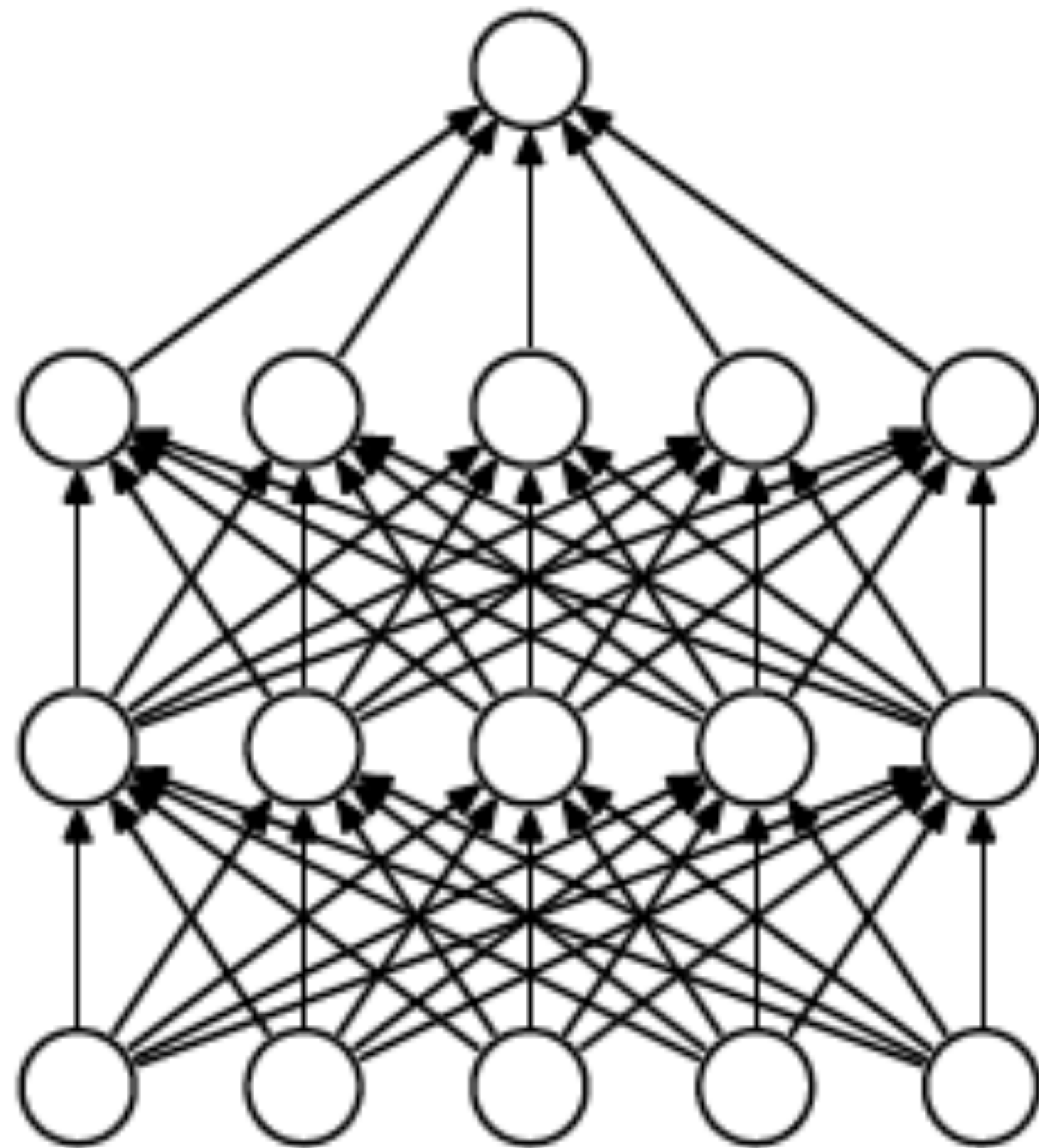
DAE = 輸送写像 (Sonoda+ 2017)

ドロップアウト (Srivastava+ 2014)

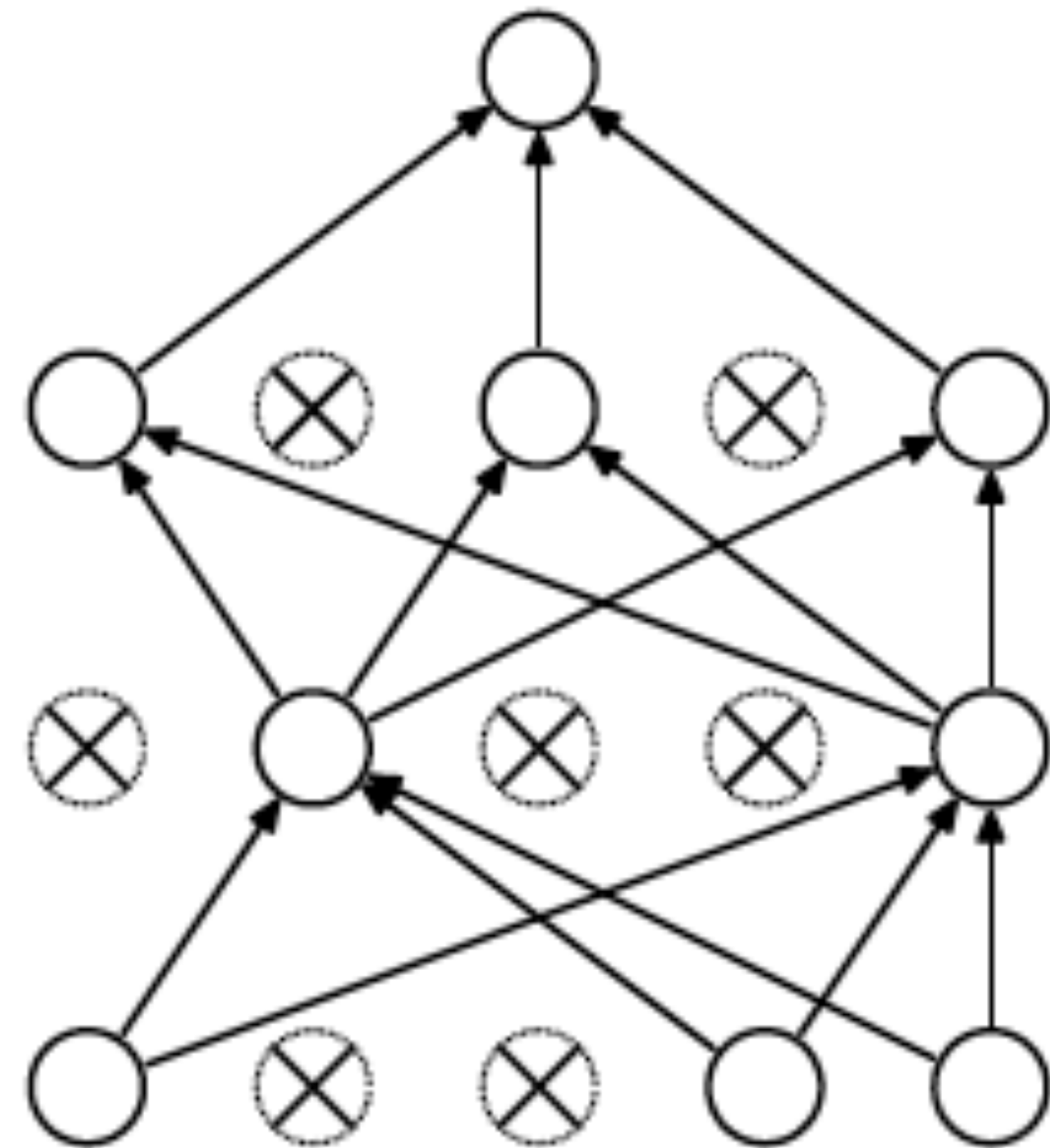
ショートカット

Residual Network (He+ 2015)

ドロップアウトは過学習を防ぐ



(a) Standard Neural Net



(b) After applying dropout.

Cross Validation-EM algorithmと類似

EMアルゴリズムは局所解が多い。

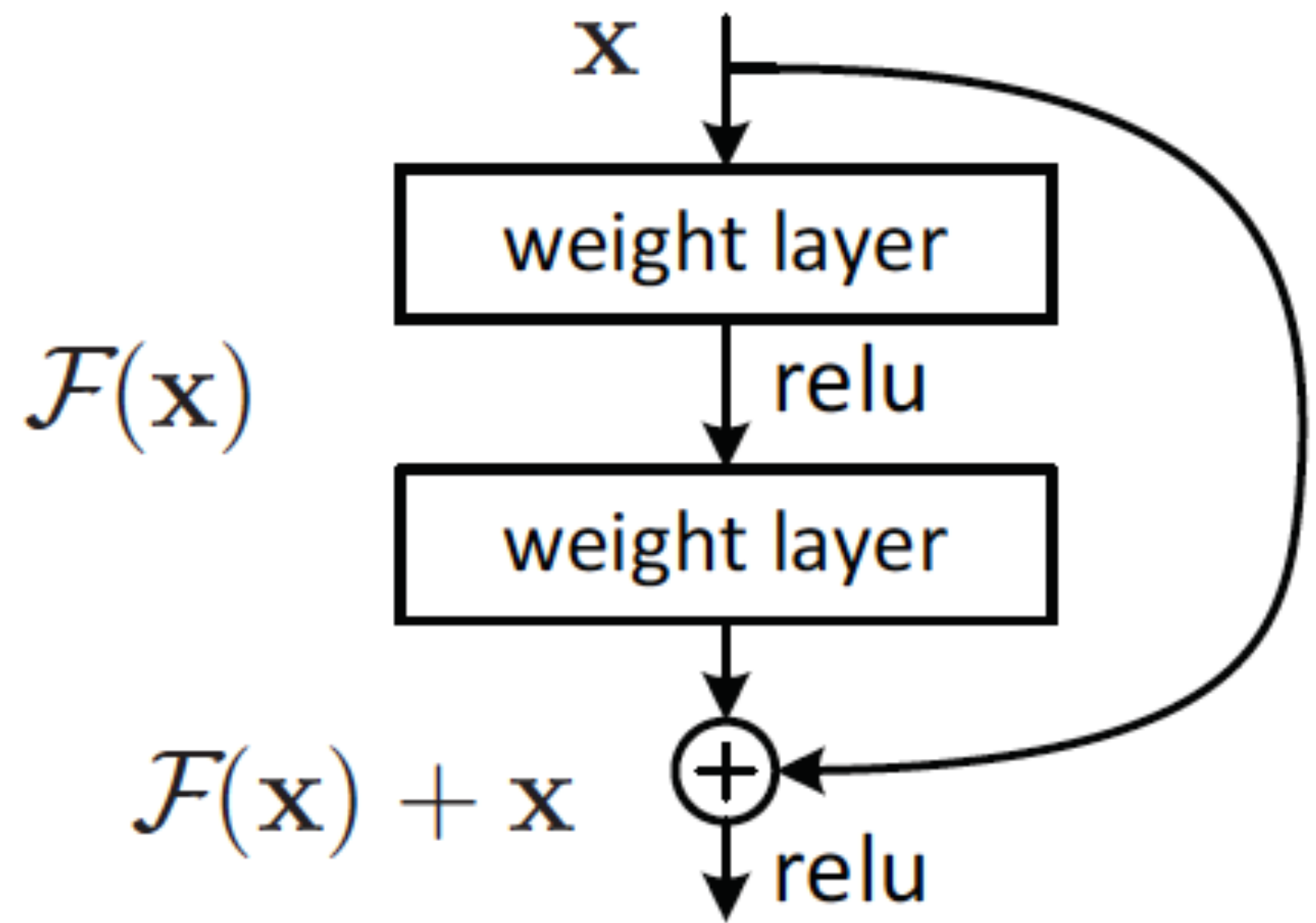
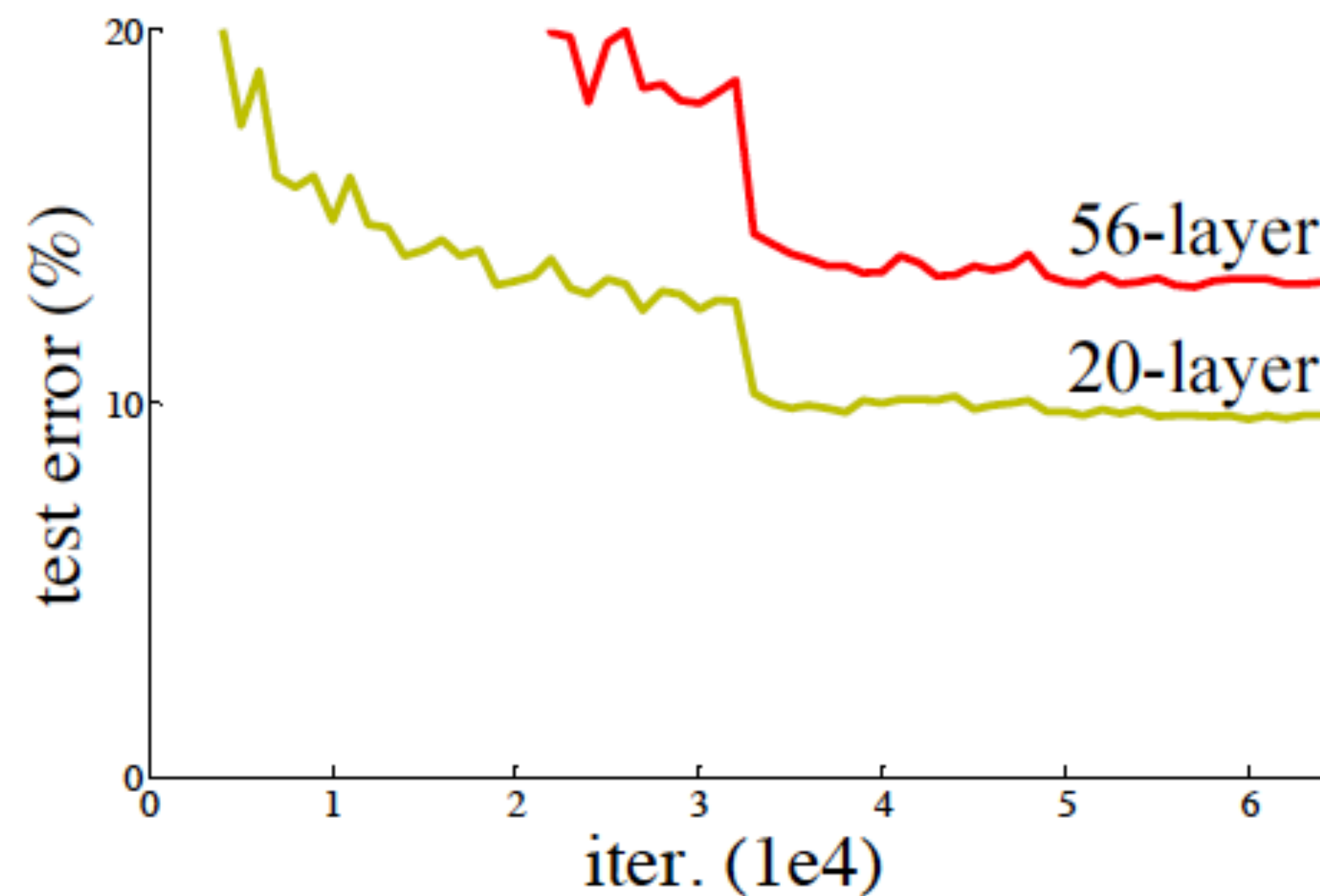
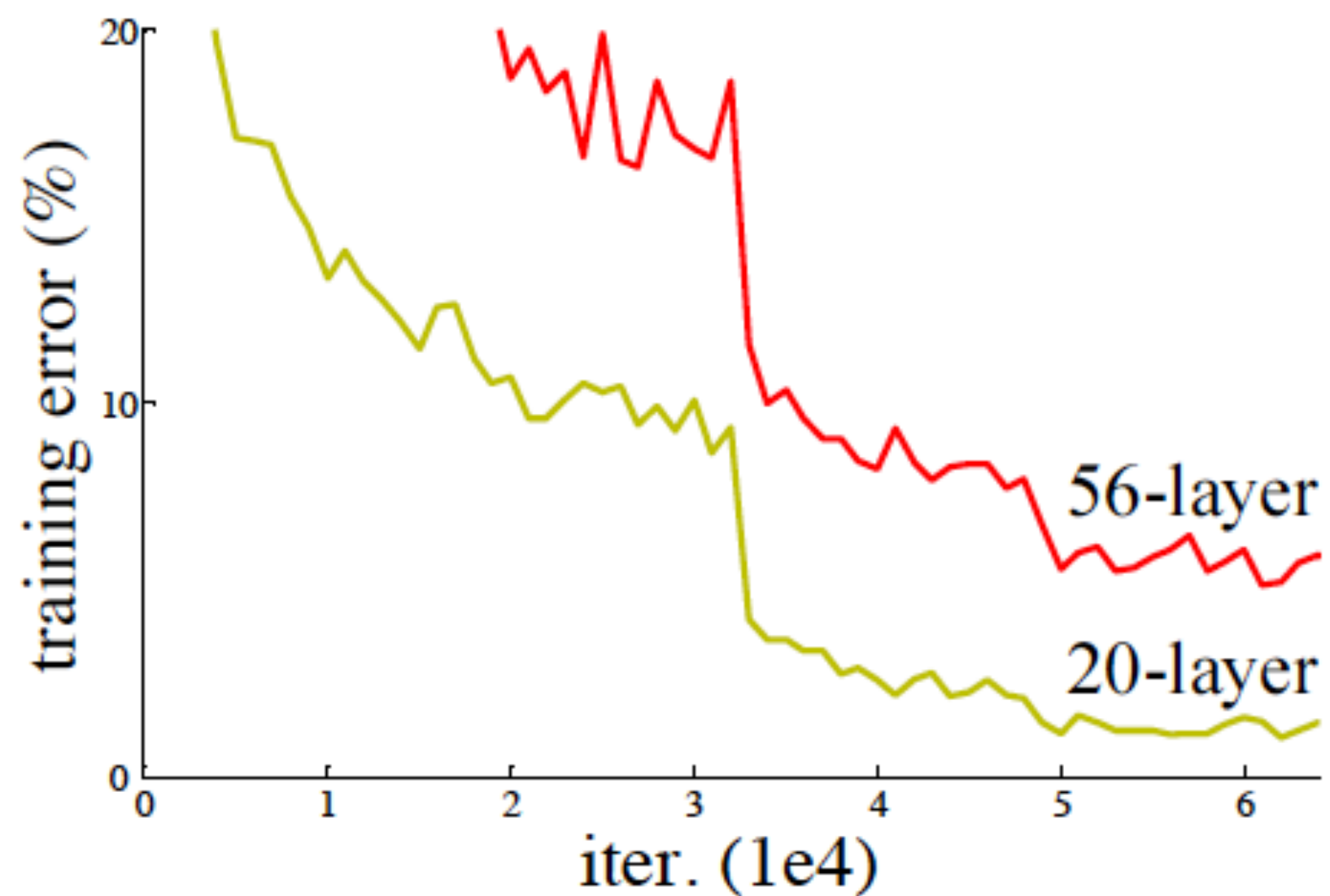
CV-EMは一部のデータを取り除く (cross-validation に類似)

CV-EMは理論解析済み (Takenouchi & Ikeda, 2010)

ショートカットは学習の停滞を防ぐ

深層学習は過学習すらしない。

Residual Network がその解決法



深層学習は無限化したモデルの離散化

ニューラルネットの積分表現理論を用いて ResNet を解析

$$g(x) = \sum_j T_j \eta(a_j \cdot x - b_j)$$

$$g(x) = \int T(a, b) \eta(a \cdot x - b) d\mu(a, b)$$

理論がアルゴリズムを向上させる

パラメータ設定の指針

性能の事前評価

必要なデータ量の目安

深層学習を凌駕する「Beyond Deep Learning」への第一歩