

TOSHIBA 革新的生産性を実現する自律学習AI (自ら学ぶAI)

①「技術」で実現を目指すCPSテクノロジー企業像

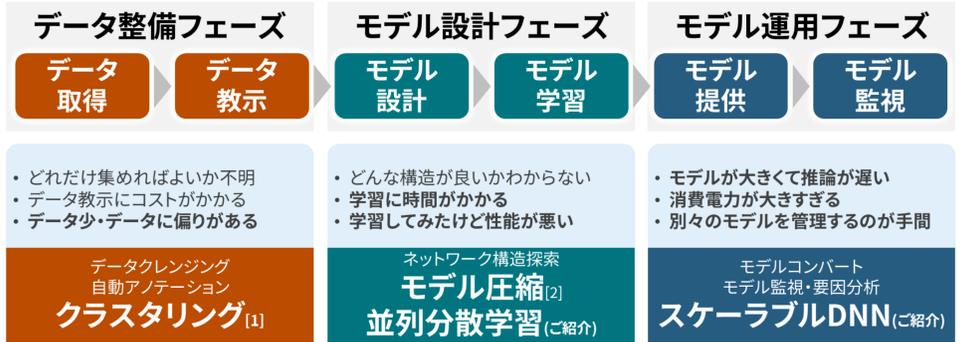
CPSテクノロジー企業の柱となるAI技術を深耕、現在12のIoTサービスを実装中



②自ら学ぶAIを実現するために

深層学習を”ビジネスの現場”で活用していくためには様々な課題がある

AIモデル開発のPLM (Product Lifecycle Management)



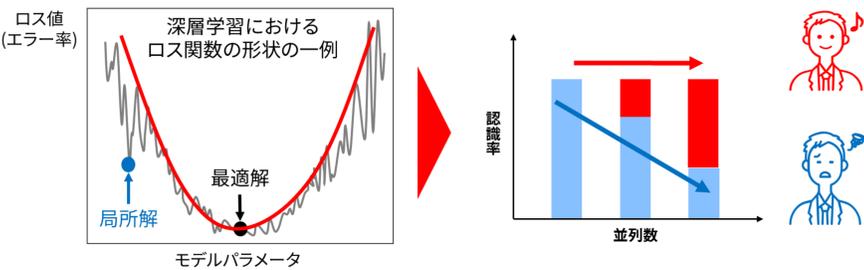
[1] Y. Tao, et al., "RDEC: Integrating Regularization into Deep Embedded Clustering for Imbalanced Datasets," In ACM 2018. <https://arxiv.org/abs/1812.02293>.
 [2] A. Yaguchi et al., "Adam Induces Implicit Weight Sparsity in Rectifier Neural Networks," In ICMLA 2018. <https://arxiv.org/abs/1812.08119>.

③大規模並列分散学習における課題

並列分散学習

大規模な並列分散学習は並列数増加に伴い精度が劣化

ミニバッチ勾配の分散が減少し、局所解に陥りやすくなるのが原因



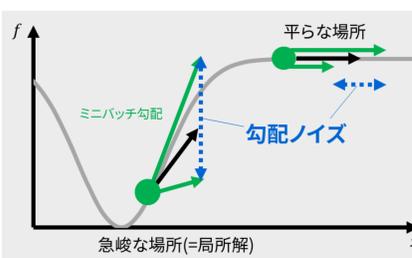
ロス関数を平滑化することで局所解に落ち込む問題を回避
→勾配ノイズを活用することで、適応的な平滑化を実現

④大規模並列時に認識精度を維持できる並列分散学習技術

並列分散学習

1 勾配ノイズはロス関数の急峻な場所(=局所解)を適応的に検知

勾配ノイズは局所解で分散が大きくなる



勾配ノイズによる適応的・低コストなロス関数平滑化を実現

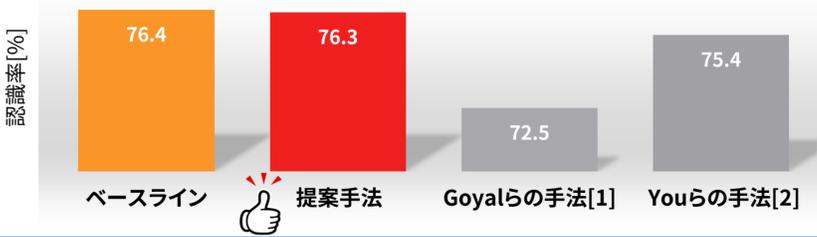
[1] K. Haruki et al., "Gradient Noise Convolution (GNC): Smoothing Loss Function for Distributed Large-Batch SGD," <https://arxiv.org/abs/1906.10822>.

⑤公開データセットにおける性能ベンチマーク

並列分散学習

公開データセット 一般物体認識(1000クラス分類): ImageNet (1024GPU並列、90epoch学習)
比較対象のネットワークはResNet-50

評価方法 (1) ベースライン: 8GPU並列で学習
(2) Goyalらの手法^[1]: バッチサイズに応じた学習率を適用する従来手法
(3) Youらの手法^[2]: レイヤーごとに学習率を制御する従来手法



大規模並列分散学習において世界最高性能を達成

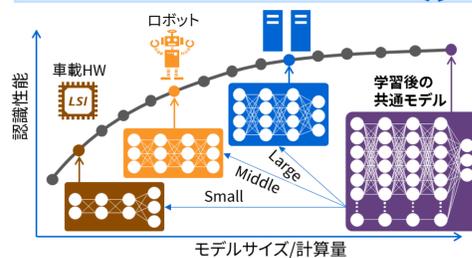
[1] P. Goyal et al., "Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour," <https://arxiv.org/abs/1706.02677>.
 [2] Y. You et al., "ImageNet Training in Minutes" <https://arxiv.org/abs/1709.05011v1>.

⑥学習後にモデルサイズや計算量を変更可能なスケーラブルDNN

スケーラブルDNN

1 スケーラブルDNNのコンセプト

学習済みの重み行列を低ランク行列に分解
自由なサイズに分解して推論可能



学習後の共通モデルを異なるアプリケーション毎に切り出して利用可能

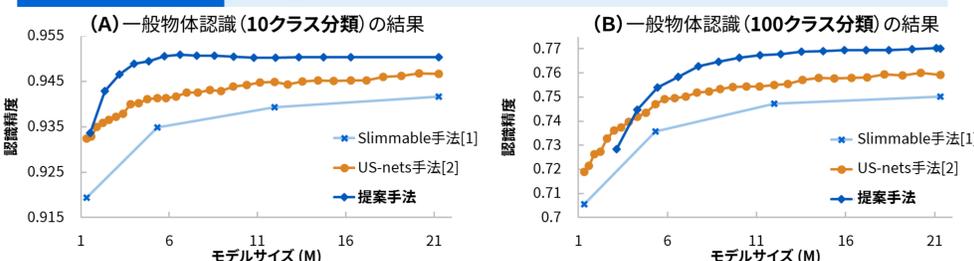
[1] A. Yaguchi et al., "Scalable Deep Neural Networks via Low-Rank Matrix Factorization," <https://arxiv.org/abs/1910.13141>.

⑦公開データセットにおける性能ベンチマーク

スケーラブルDNN

公開データセット (A) 一般物体認識(10クラス分類): CIFAR-10
(B) 一般物体認識(100クラス分類): CIFAR-100

評価方法 (1) Slimmable手法^[1]: 予め定義したモデルサイズに分解可能な従来手法
(2) US-nets手法^[2]: 任意のモデルサイズで分解可能な従来手法
モデル構造はResNet-34



広いモデルサイズ帯で従来研究を上回る高い認識精度を達成

[1] J. Yu et al., "Slimmable Neural Networks," In ICLR 2019. <https://arxiv.org/abs/1812.08928>.
 [2] J. Yu et al., "University Slimmable Networks," In ICCV 2019. <https://arxiv.org/abs/1903.05134>.

⑧まとめ

01 東芝が目指す姿

世界有数のCPSテクノロジー企業へ



02 AI技術開発の方向性



03 社会実装の方向性



04 AIモデル開発におけるモデル設計・運用を支援する並列分散学習、スケーラブルDNNを開発

