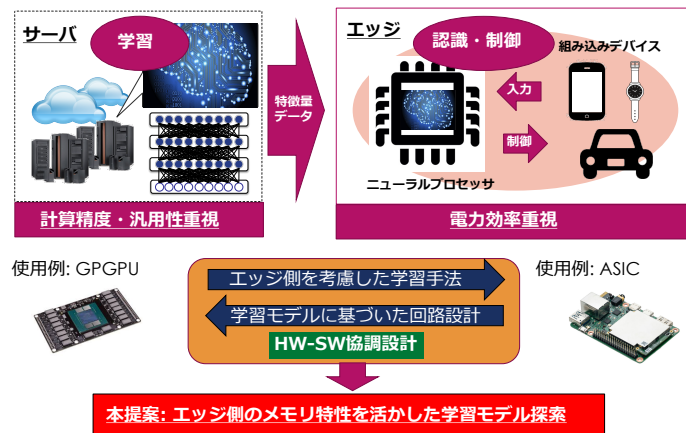
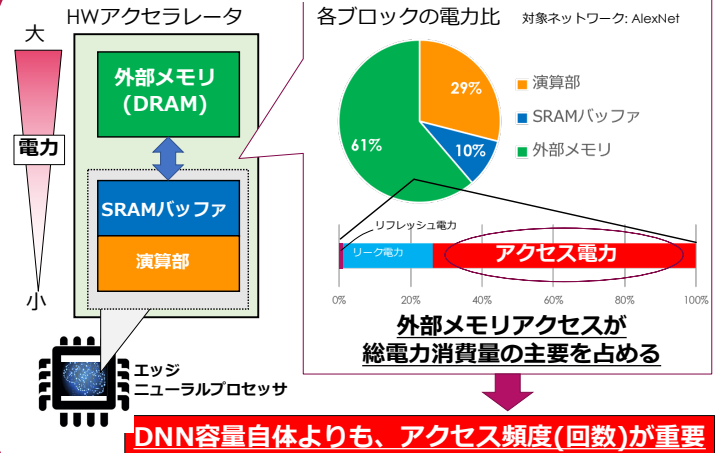


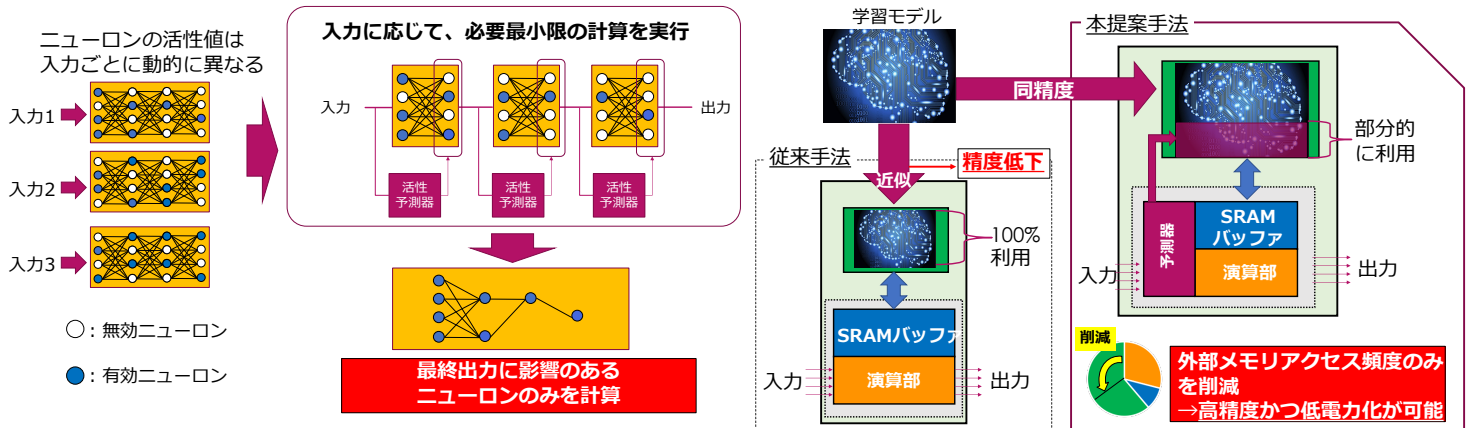
1. 低電力計算環境下でのAI計算



2. 実計算機の消費電力

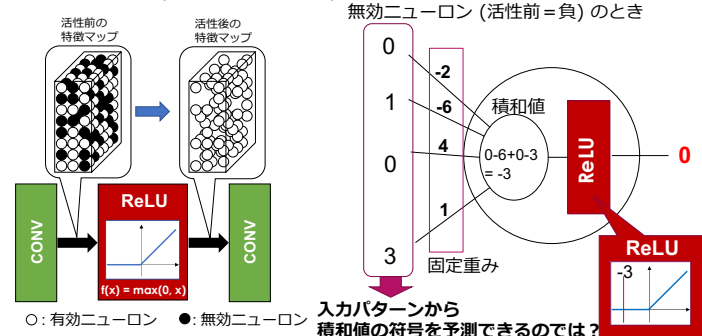


3. 提案: 動的予測による計算量・メモリアクセス数の削減

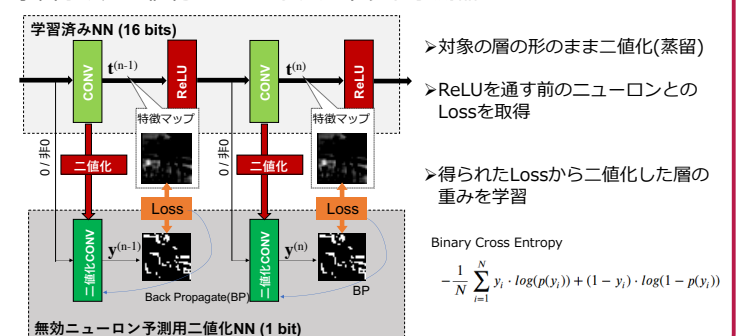


4. 手法: 無効ニューロンの予測器学習

▶ 活性化関数 ReLU (Rectified Linear Unit)

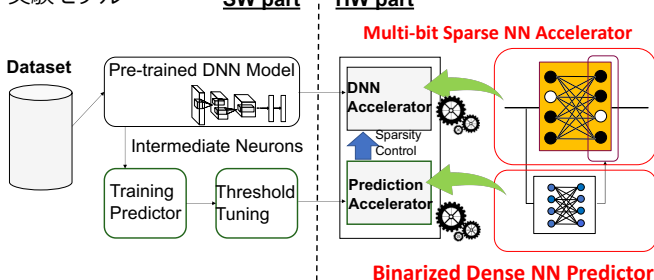


学習手法: 二値化NNによるサブネット予測器



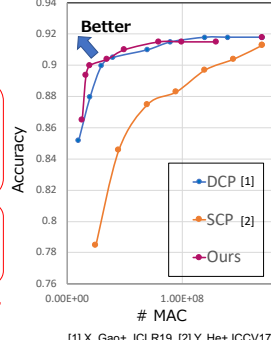
5. 結果: 計算量と電力量の削減

実験モデル

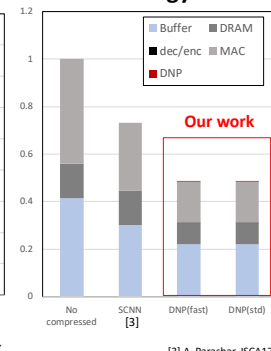


実験結果

VGG for CIFAR10



Energy



まとめ

- ▶ DNN計算特性を活かした、ハードウェア考慮モデルの提案
- ▶ 計算量と電力の削減を比較評価

今後の課題

- ▶ 精度 vs. 電力を統一的に設計するフレームワーク
- ▶ 性能(速度)に対する効率化
- ▶ アプリケーション特性を活かしたモデルの設計