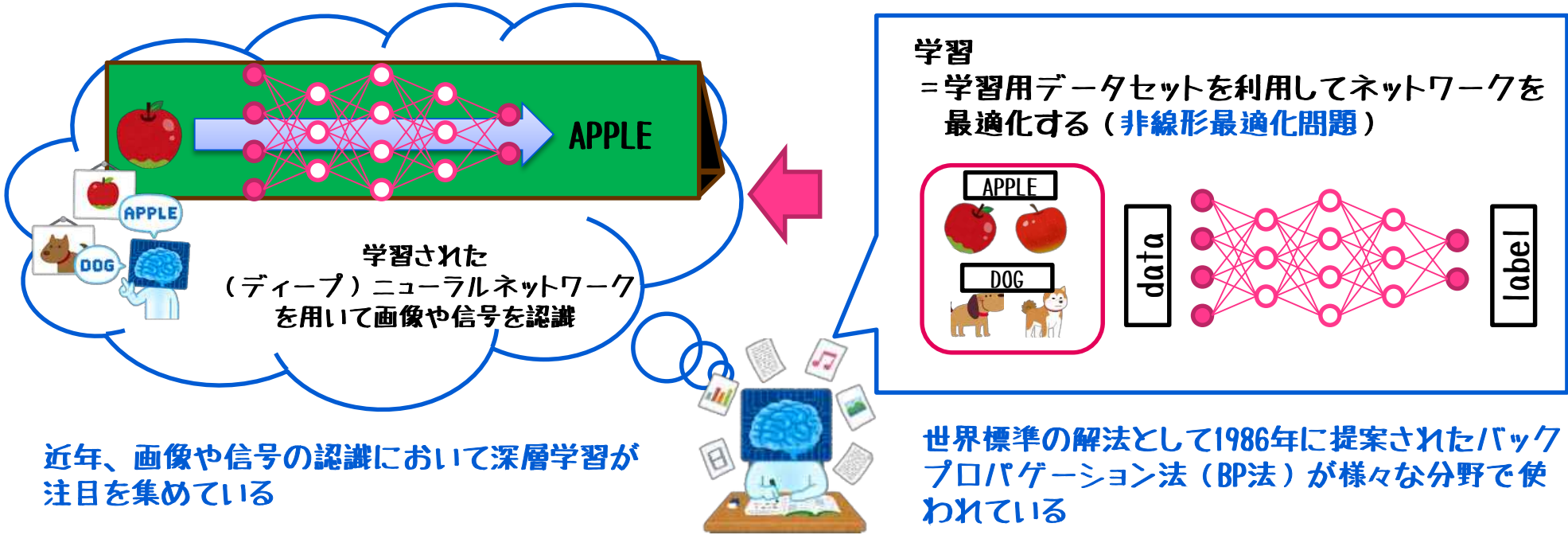


新しい深層学習アルゴリズム

研究背景・解決すべき問題



解決方法・波及効果

成果

- 新しい深層学習アルゴリズムを開発
- 近年のスーパーコンピュータと親和性の高い、行列分解を基盤とする
 - 高速！ 高並列性！！
- 「スーパーコンピュータ x 深層学習」によるサイバーフィジカル社会の実現へ！！

- 「非線形非負行列分解」を利用した新しい深層学習アルゴリズムを開発
- 古典的なBP法と比較して
 - 同程度の認識性能を実現 (手書き文字認識で98.5%程度の認識率)
 - 高速・高並列性を実現

提案法の基本アイデア

対象：全結合DNNモデル

- ✓ DNNモデル：活性化関数：ReLU
- ✓ 目的関数：二乗誤差 + 正則化項

$$f(x) = \max(0, x)$$

$$\min_{W,b} E + \lambda \sum_{i=1}^d \|W_i\|$$

$$E = \|Y - W_d f(W_{d-1} \dots f(W_1 X + b_1 \mathbf{1}^T) \dots + b_{d-1} \mathbf{1}^T) + b_d \mathbf{1}^T\|$$

アルゴリズムの基本アイデア

- ✓ 最適化ステップ：非線形非負値分解に基づく逐次最適化
 - ・非負行列分解 (semi-NMF)

$$\min_{W_d, b_d, (Z_{d-1} \geq 0)} \|Y - (W_d Z_{d-1} + b_d \mathbf{1}^T)\|_F + \alpha \|W_d\|$$

・非線形semi-NMF

$$\min_{W_i, b_i, (Z_{i-1} \geq 0)} \|\hat{Z}_i - f(W_i Z_{i-1} + b_i \mathbf{1}^T)\|_F + \alpha \|W_i\| + \beta \|Z_{i-1}\|, \\ i = d-1, d-2, \dots, 2$$

・非線形最小二乗問題

$$\because Z_i = f(W_i Z_{i-1}) \geq 0$$

$$\min_{W_1, b_1} \|\hat{Z}_1 - f(W_1 X + b_1 \mathbf{1}^T)\|_F + \alpha \|W_1\|$$

- ✓ 更新ステップ：計算した W から Z を更新
- ✓ 反復計算：得られたW,Zを初期値として最適化・更新ステップを繰り返す

数値実験・結果

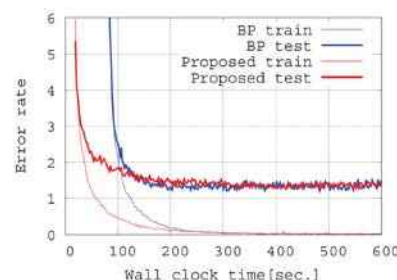
数値実験1：認識性能

- ✓ テスト問題：画像認識 (MNIST)
 - ・手書き数字認識
 - ・画像サイズ：28 x 28
 - ・学習データ：60,000枚
 - ・テストデータ：10,000枚



- ✓ パラメータ
 - ・隠れ層：1000-500

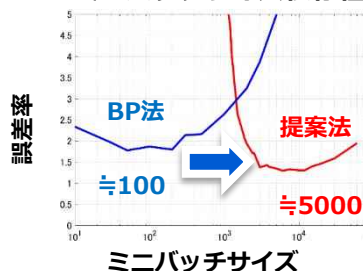
結果1： 古典的なBP法と同程度の認識性能を達成



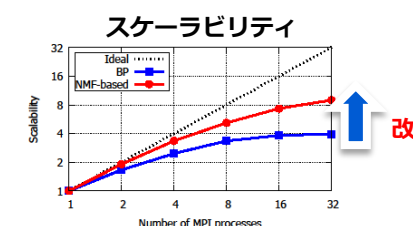
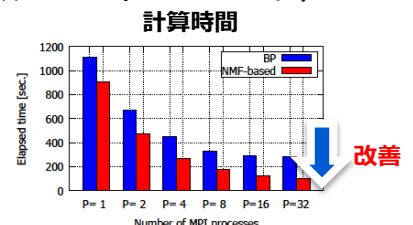
数値実験2：並列性能

- ✓ テスト問題：画像認識 (MNIST)
- ✓ 計算環境：京コンピュータ@理研AICS (1-64ノード)

BP法・提案法の性能のミニバッチサイズ依存性



これに基づく
並列実装法



結果2： 高速・高並列性を実現