

BPA-PR法による 階層型ニューラルネットワークの故障補償能力の向上

山森一人¹⁾森元聡明²⁾吉原郁夫³⁾

Defect Compensation method for Multi-Layer Neural Network by PR Scheme with BPA algorithm

Kunihito YAMAMORI

Akira MORIMOTO

Ikuro YOSHIHARA

Abstract

Some researchers have been proposed to implement neural networks into Wafer Scale Integration(WSI) to achieve fast learning. When neural networks are implemented into a WSI, it has to have a mechanism to avoid hardware defects. To compensate hardware defects, the partial retraining (PR) scheme has proposed. The performance of PR scheme depends on the weights in the neural network because PR scheme only adjusts the weights belonging to a neuron affected by the defects. In this paper, we propose back propagation with annealing scheme (BPA scheme) to improve defect compensation ratio. We show that BPA scheme achieved higher capability of defect compensation than that of conventional BP algorithm.

Key Words:

Back Propagation with annealing, Partial Retraining scheme, Multi-Layer Neural Network, Fault Tolerance

1 はじめに

近年、工学上の応用技術の1つとしてニューラルネットワーク(NN:Neural Network)が広く使われるようになった。一方、NNには学習に時間を要するという問題があり、大規模問題への応用が難しい。そこで、近年のVLSI技術の発展を踏まえ、学習高速化のためNNをWSI(Wafer Scale Integration)で実現する研究が行われている。しかし、多数のニューロンが必要な大規模システムをWSIで実現する場合、WSI中心部ではかなりの高温になることが予想され、これにより故障の増大が危惧されている。したがって、WSIではシステム製造時の欠陥のみならず製造後の故障への対策も不可欠である。

WSIへ実装したNN上の故障を回避する手法は、数多く提案されている。高田ら¹⁾は、縮退故障(出力が0または1に固定される故障)について、学習パターンにノイズを混入する耐故障化手法を提案している。山森ら²⁾は、故障の影響を受けるニューロンのみを対象として、誤差逆伝搬学習(BP)法を適用することにより故障補償を行う部分再学習(PR)法について提案している。伊藤ら³⁾は、既存の耐故障化手法を拡張し、リンクの故障をニューロン素子出力の0縮退故障に置き換えて学習を行うという手法を提案し、これにより少ない学習時間で高い耐故障性を実現できることを示した。また、ハードウェアの観点からは、菅原ら⁴⁾が、予備ニューロンを用いた故障補償手法を提案し、その回路設計を行っている。

本論文では、これらの中でも本質的に故障検出のための特別な回路が不要であり、製造後の任意の故障に迅速に対処できるPR法に着目する。

¹⁾情報システム工学科准教授

²⁾情報システム工学科学生(現在(株)富士通ピー・エス・シー)

³⁾情報システム工学科教授

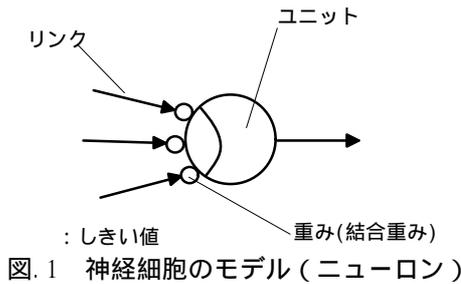


図. 1 神経細胞のモデル (ニューロン)

PR法を用いる際には、BP法などであらかじめ学習を済ませておく必要がある。このとき、特定のニューロンやリンクが学習精度の維持に必要不可欠であったとき、故障前の状態から再学習により故障補償を行うPR法では当該箇所の故障は補償することができない。すなわち、PR法で高い故障補償率を達成するためには、NN内のニューロンが等しく役割を分担するよう初期学習を行っておくことが必要である。

本研究は、部分再学習法による故障補償において、初期学習を工夫することで故障補償率を向上させることを目的とする。そのため、誤差逆伝搬学習法にシミュレーテッドアニーリング法の考え方を援用したBPA (Back Propagation with Annealing) 法を初期学習に用いることを提案し、顔画像認識問題により故障補償能力と平均学習時間について議論する。

2 ニューラルネットワーク

2.1 階層型ニューラルネットワーク

脳の神経系は神経細胞を単位として構成されており、それらが互いに結びついている。この神経細胞の仕組みをモデル化したものがニューロンであり、その概念を図1に示す。ニューロンがいくつか結合することによってNNが構成される。NNには、ニューロン間の結合が順方向のみの階層型ニューラルネットワーク (MLN) と、ユニット間が相互に結合した相互結合型ニューラルネットワークがある。本研究では、MLNについて議論する。MLNの例を図2に示す。MLNは、入力層、中間層、出力層の3種類の層からできている。中間層は1層とは限らず、階層間は重み付きのリンクで全結合されており、階層内での結合はない。

2.2 誤差逆伝搬学習法

MLNの代表的な学習法として誤差逆伝搬学習 (BP:Back Propagation) 法⁵⁾がある。BP法はNNにある入力を与えたときに得られる出力値と教師

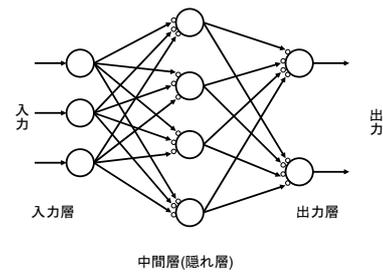


図. 2 3層階層型ニューラルネットワークの例

信号の値との二乗誤差が小さくなるよう、最急降下法により重みの値を修正する方法である。

3 故障補償法

3.1 故障の定義

本研究では、(1)式を満たしていたNNが何らかの理由により(1)式を満たさなくなることをNNの故障と定義する。

$$\max_{p,n} \sqrt{\left(t_n^{(p)} - o_n^{(p)}\right)^2} < e. \quad (1)$$

(1)式において n 、 p はそれぞれ出力ニューロン番号、学習パターン番号を表し、 $t_n^{(p)}$ は、学習パターン p の出力ニューロン n に対する教師信号、 $o_n^{(p)}$ は、学習パターン p における出力ニューロンの出力値である。

ニューロンの構造からNNの故障箇所は、一種のしきい素子であるユニット、リンク、及び重みの3つが考えられる。故障の種類にはスタック故障とドリフト故障があるが、本研究ではスタック故障を対象とする。スタック故障とは、入出力信号に関わらず値が常に一定となる故障のことである。リンクの断線故障は重みの0-スタック故障に相当する。よって本研究では、入力層-中間層間のリンクの0-スタック故障とユニットの0、1-スタック故障を対象とした。

3.2 部分再学習法

本研究では、部分再学習 (PR:Partial Retraining) 法を用いて故障補償を行う。PR法を用いて故障補償を行う際には、BP法などによりあらかじめNNを学習させておく必要があり、これを初期学習と呼ぶ。各ニューロンはメモリを持ち、入力層ニューロンのメモリには学習パターンの入力値、中間層ニューロンのメモリには初期学習後の各学習パターンに対する各中間ニューロンの出力値及び入力層-中間層間の重みの値、出力層ニューロン

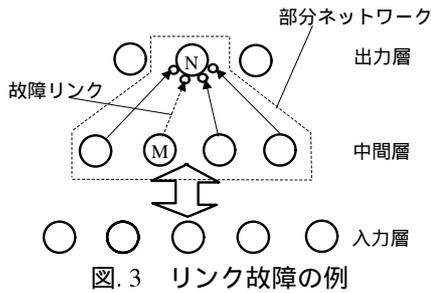


図.3 リンク故障の例

のメモリには学習パターンの教師信号と中間層-出力層間の重みの値が蓄えられ、これらは故障発生時の再学習に用いられる。

図3のように中間層のニューロンMと出力層のニューロンNの間のリンクが故障を起こした場合を考える。このとき故障の影響を受けるのは、図3からわかるように出力層のニューロンNだけである。そこで、ニューロンNを出力ユニットとする2層の部分ネットワークにおいて、BP法を用いることで故障補償を行う。このとき、入力信号として中間層のニューロンが持つ故障発生前の出力値を用いる。

3.3 BPA法

初期学習において、特定のニューロンがある学習パターンの記憶に不可欠な役割を持つよう学習が進んだ場合、このニューロンが故障してしまうとPR法では故障を補償できない。したがって、複数のニューロンが分担して学習パターンを記憶するよう初期学習させることが望ましい。そこで、BP法における重み修正時に人為的にノイズを加えるBPA (Back Propagation with Annealing) 法を提案する。重み修正時にノイズを加えることで、特定のニューロンが特定の学習パターンを記憶するよう学習が収束するのを妨げる。すなわち、BPA法はBP法のように最急降下法により最適解を探索するのではなく、最急降下を敢えて阻害することでニューロンの役割分担を狙ったものである。

BPA法におけるノイズ値の与え方にはシミュレーテッドアニーリング (SA) 法⁶⁾を援用する。SA法とは、組合せ最適化問題でローカルミニマムに捕まらず、グローバルミニマムを探索可能にするアルゴリズムである。SA法での温度パラメータ T の変遷に従ったグローバルミニマム探索範囲の変化を、BPA法では学習回数(温度パラメータ)の増加につれてノイズ値を小さくすることで対応させる。最終的には学習を収束させる必要が

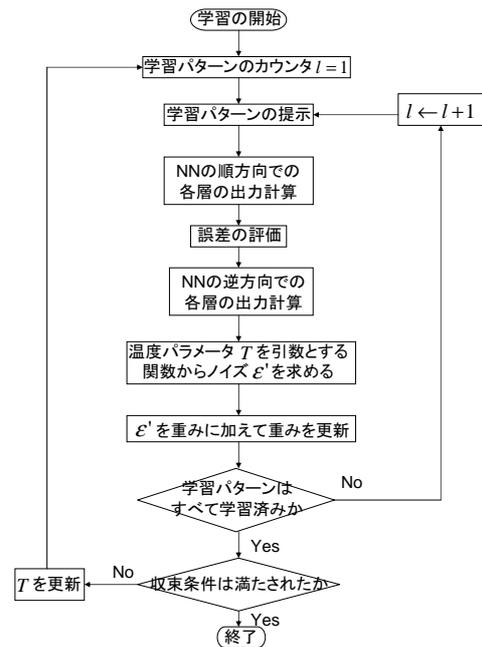


図.4 BPA法のアルゴリズム

あることから、BPA法で加えるノイズ項は学習の進行に伴って単調減少しなければならない。そこで(2)式に示したノイズ $\varepsilon' (\geq 0)$ を重みの更新量に加えた(3)式により、初期学習時の重み更新量を求める。

$$\varepsilon' = \left(1 - \frac{T}{T_0}\right) \times 0.00001, \quad (2)$$

$$\Delta w_{k-1,m}^{k,n}(p+1) = w_{k-1,n}^{k,n}(p) + \Delta w_{k-1,m}^{k,n} + \varepsilon'. \quad (3)$$

ここで、 T_0 は定数、 T は学習のエポック数であり、 $\Delta w_{k-1,m}^{k,n}(p)$ は、学習パターン p が提示されたときの $(k-1)$ 層の第 m ニューロンから k 層の第 n ニューロンへのBP法で求めた重みの更新量である。図4にBPA法のアルゴリズムをフローチャートとして示す。

4 故障補償実験

4.1 実験条件

顔画像認識問題を用いて故障補償能力を評価する。256×256画素の顔画像32枚を用意し、それぞれを8×8のブロックに分割後、各ブロックの平均輝度値を正規化して入力信号とする。顔画像1枚あたり1つの出力ニューロンを用意し、対応する顔画像が入力された時に0.9、それ以外の画像が入力された時は、0.1を出力するように学習

表.1 学習パラメータ

| | |
|--------------------|----------|
| ネットワークサイズ | 64-15-32 |
| 学習率 | 0.70 |
| 慣性項 | 0.40 |
| 最大学習回数 | 10000 |
| 学習終了判定誤差 | 0.17 |
| 実験回数 | 90 |
| T の初期値 (T_0) | 400~1200 |

を行った。NNの構成は、入力層に64個、中間層に15個、出力層に32個のニューロンを持つMLNである。故障補償性能は、ある故障パターンを想定したときに全ての学習パターンで(1)式が成立した場合のみを故障補償成功としてカウントし、これを故障箇所を変えて故障補償実験を行った回数に対する割合として算出した故障補償率で評価する。故障補償率の定義を式(4)に示す。

$$\text{故障補償率 (\%)} = \frac{N_{\text{compensated}}}{N_{\text{all}}} \times 100. \quad (4)$$

ここで、 $N_{\text{compensated}}$ は故障補償後にすべての学習パターンにおいて(1)式を満たすことができたNNの数、 N_{all} は実験回数である。重みの値は乱数で初期化しているので、重みの初期値によって故障補償率が変化する可能性がある。よって各故障パターンに対して重みの初期値を変化させて3回の実験を行う。したがって、 N_{all} は30の故障パターンに対してそれぞれ3回実験を行うので $30 \times 3 = 90$ となる。

また、リンク故障とユニット故障において故障補償時間も測定する。90回の故障補償実験においてそれぞれ時間を測定し、その平均を求める。本実験で用いたパラメータを表1に示す。 T_0 はその値を変化させて実験的に適切な値を決定した。

4.2 実験結果

4.2.1 リンクの故障補償実験

リンクの0-スタック故障に対し、初期学習にBP法を用いてPR法を適用したとき(BP+PR)と、初期学習に様々な T_0 でBPA法を用いた後にPR法を適用したとき(BPA+PR)の故障補償率を図5に示す。 T_0 を800としたBPA+PRでの故障補償率は96.6%となる一方、BP+PRでは73.3%の故障補償率であった。したがって、適切な T_0 を用いたBPA+PRを用いることで故障補償率を向上させることが可能なことが分かった。

BPA法の効果を検証するため、より解析が容易なXOR問題を用いて調査を行った。図6にXOR

問題をBPA法とBP法でそれぞれ学習させた後、(0,1)を入力したときの中間ニューロンの出力を示す。図6からBP法で初期学習を行った場合、1番や4番のニューロンが故障すると、値が大きいため回復不可能の恐れがある。なぜならこれらのニューロンは比較的大きな出力であり、これが失われたときの影響が大きいためである。一方、BPA法で初期学習を行うとこれらのニューロン出力の値がBP法よりもやや低い値となると同時に、5番、6番ニューロンの出力が大きくなっており、役割の分担が行われたと考えられる。実際に4番ニューロンを故障させて故障補償実験を行ったところBP法を初期学習としたときは回復できなかったが、BPA法を初期学習としたときは回復させることが出来た。一方、1番ニューロンが故障した場合はどちらの手法でも補償できなかった。

図7、図8にPR法適用前と後の中間層ニューロンの出力に重みを掛けた値を示す。故障は第4ニューロンの0-スタック故障としている。第4ニューロンの故障によってPR法適用後の値は0となっている。第4ニューロンの値が0となったことで、PR法により全ニューロンに対して重みの修正が行われ出力層からの出力を維持しようとしていることがわかる。

平均故障補償時間についても、BP+PRよりもBPA+PRの方がより短い時間で学習が終了した。図9に示した平均故障補償時間から、BPA+PRは0.124秒、BP+PRは0.14秒の平均故障補償学習時間であったことが分かる。BPA+PRでは、ノイズを重み修正時に加えたことで初期学習の時間がBP法より増大する一方、PR法で故障補償に要する時間が短縮されるため、トータルでの処理時間も短縮したものと考えられる。

4.2.2 ユニットの故障補償実験

ユニットの0-スタック故障での故障補償率を図10、1-スタック故障での故障補償率を図11に示す。また、それぞれの故障での平均故障補償時間を図12、図13に示す。故障補償率では、リンク故障補償時と同様、BP+PRよりBPA+PRの方が向上した。0-スタック故障ではBPA+PRで96.6%、BP+PRで73.3%の故障補償率、1-スタック故障では、BPA+PRで93.3%、BP+PRで80.0%の故障補償率であった。

平均故障補償時間は、0-スタック故障に関してはBPA+PRが0.118秒に対しBP+PRでは0.155

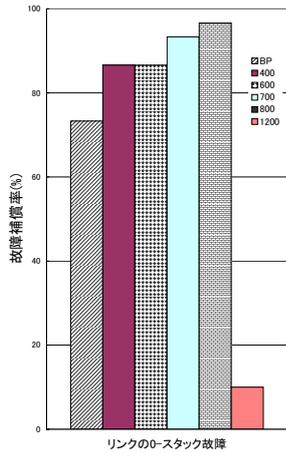


図. 5 T_0 を変化させたときのリンク故障時の故障補償率

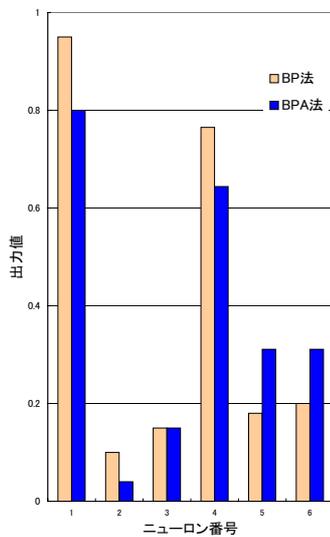


図. 6 XOR問題の入力(0,1)でのBPA法とBP法の間層ニューロンの出力

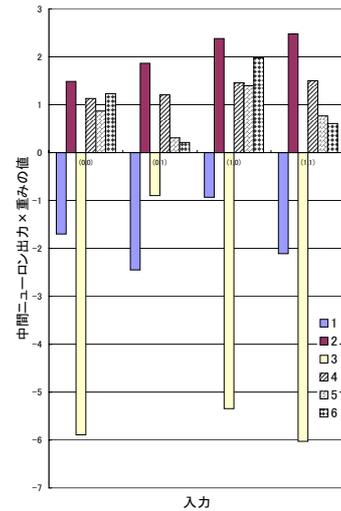


図. 7 XOR問題でのPR法適用前の中層ニューロンの出力×重みの値

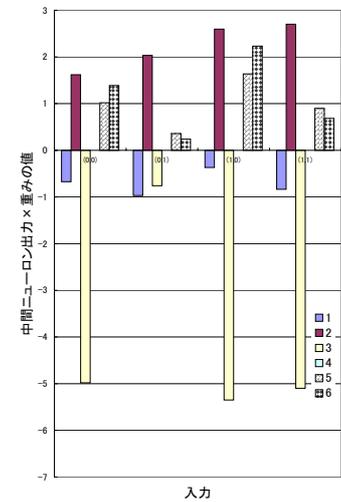


図. 8 XOR問題でのPR法適用後の中層ニューロンの出力×重みの値

秒であり、BPA+PR法がより短い時間で故障補償することができた。しかし、1-スタック故障に関してはBP+PRが0.163秒、BPA+PRが0.142秒となり、BPA+PRの方が増加する結果となった。この理由は、重み修正時に加えたノイズ分により、誤差を十分小さくするため初期学習回数が増加し、PR法適用時の学習時間短縮を考慮してもトータルの学習時間が増加したためと考えられる。

5 おわりに

本研究では、WSI実装されたMLNの故障補償手法の1つであるPR法を用いたハードウェアNN上での故障補償において、PR法とBPA法を組み合わせることで故障補償率の向上を目指した。実験の結果、故障補償率については、リンクの0-スタック故障、ユニットの0、1-スタック故障のいずれも初期学習にBPA法を用いることでBP法

と比べて故障補償率の向上が見られた。しかし、平均故障補償時間を比較すると、ユニットの1-スタック故障でBPA+PRの方がBP+PRに比べて増加する場合があることが分かった。

今後の課題は、大規模問題での実験、他の故障でのBPA法の検証と評価などが挙げられる。

参考文献

- [1] 高田広之、高浪五男、“階層型ニューラルネットワークの耐故障化 -ニューロン出力の縮退故障-”、機能集積情報システム研究会発表論文集Ⅲ、pp.100-107、2000。
- [2] 山森一人、阿部亨、堀口進、“部分再学習法による故障補償を行ったニューラルネットワークの汎化能力”、機能集積情報システム研究会発表論文集Ⅲ、pp.108-115、1998。
- [3] 伊藤岳広、高浪五男、“階層型ニューラルネットワークの多重故障に対する耐故障性”、機能集積情報

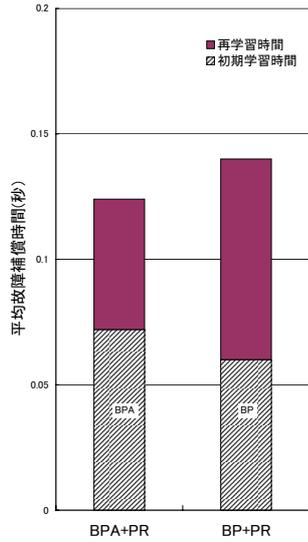


図. 9 リンク故障時の平均故障補償時間

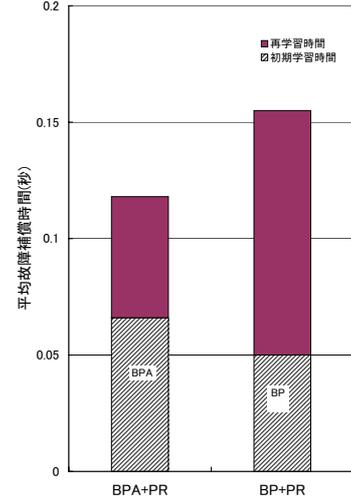


図. 12 ユニットの0-スタック故障時の平均故障補償時間

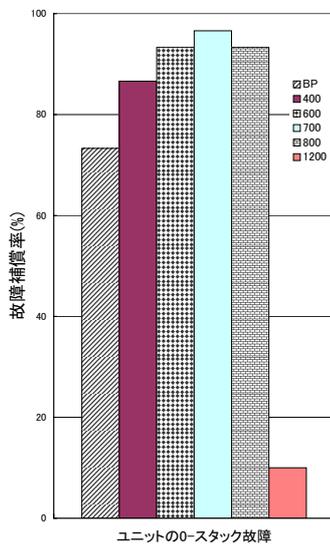


図. 10 ユニットの0-スタック故障時の故障補償率

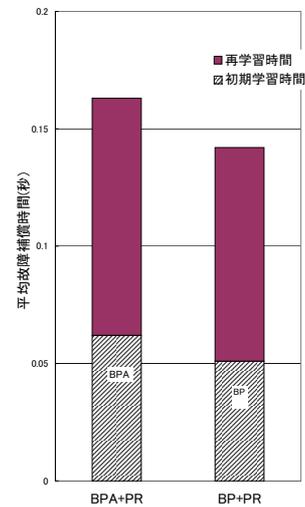


図. 13 ユニットの1-スタック故障時の平均故障補償時間

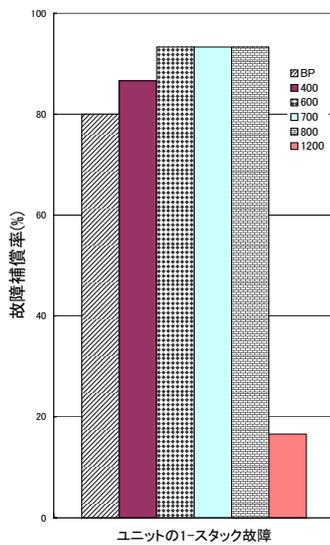


図. 11 ユニットの1-スタック故障時の故障補償率

報システム研究会発表論文集、pp.252-263、1997.

- [4] 菅原英子、堀口進、"FPGA による階層型ニューラルネットワーク故障補償のハードウェア化"、機能集積情報システム研究会発表論文集 III、pp.238-244、2001.
- [5] 熊沢 逸夫、"学習とニューラルネットワーク"、森北出版社、1998.
- [6] 戸谷智之、石川幹人、荒木均、星田昌紀、"シミュレーテッドアニーリングによるアミノ酸配列解析"、情報処理学会創立45周年記念DVD、1993