

平成 22 年 6 月 18 日現在

研究種目：若手研究 (B)

研究期間：2007～2009

課題番号：19700226

研究課題名 (和文) 分割階層型 CNN と進化的構造適応学習法による顔表情識別手法の提案

研究課題名 (英文) Proposal of the facial expression classification method based on the divided hierarchical CNN and the evolutionary structural adaptive learning algorithm

研究代表者 大枝真一 (OEDA SHINICHI)

木更津工業高等専門学校・情報工学科・講師

研究者番号：80390417

研究成果の概要 (和文)：

GA および IA を用いた進化的構造適応学習法により, 砂時計型ニューラルネットワークの最適化を行った. その結果, ニューラルネットワークの初期結合荷重をランダムに与える場合よりも, 収束速度が高速な初期結合荷重の探索に成功した. ネットワーク・インバージョンの局所解問題を解決し, それを砂時計型ニューラルネットワークに適用した DoubleBP を開発した. 提案した有効性を検証するために, 進化的構造適応学習法を分割階層型 CNN に適用し, 顔表情の識別を行った.

研究成果の概要 (英文)：

In this research, we optimized the sandglass type neural network by the evolutionary structural adaptive learning algorithm with GA or IA. As a result, we succeeded in the search of the initial connection weight in neural network. A convergence speed of the optimized weight was faster than the initial weight given to it at random. We solved the local minimal problem of the network inversion. It was applied to the sandglass type neural network. To verify the effectiveness of our proposed method, we applied the evolutionary structural adaptive learning algorithm to the divided hierarchical CNN and we classified the facial expression image data.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
19 年度	1400000	0	1400000
20 年度	500000	150000	650000
21 年度	500000	150000	650000
年度			
年度			
総計			2700000

研究分野：総合領域

科研費の分科・細目：情報学・感性情報学, ソフトコンピューティング

キーワード：ニューラルネットワーク

科学研究費補助金研究成果報告書

1. 研究開始当初の背景

(1) セラニューラルネットワーク (Cellular Neural Network, 以下, CNN とする) は, セルと呼ばれる単純なアナログ回路を格子状に配置したネットワーク構造となるものである. CNN は学習を必要とせず, 既存知識に基づいたニューラルネットワークの連想記憶能力によって推論出力を得る. また, 単純なセルの組み合わせから生成する複雑な時空間ダイナミクスを利用する非線形回路であり, 理論的な性質や画像処理への応用が盛んに進められている.

(2) CNN の各セルの状態は近傍範囲内のセルの状態の影響を受けながら微分方程式系で変化し, システムの漸近安定平衡点に収束する特性を持っている. この特性は自己想起機能といえ, CNN を連想記憶媒体として識別問題に応用する際に重要な役割を果たしている. しかしながら, 識別対象となる画像サイズが大規模になった場合, CNN を構成するセル数も多くなる. そのため, 漸近安定平衡点までの収束には多大な計算量を必要とするとともに, 識別能力も低減することが考えられる.

2. 研究の目的

(1) 研究背景に基づき, 著者らは分割階層型 CNN の研究に着手した. この提案手法では, 識別対象となる入力信号パターンをいくつかの領域に分割し, それぞれの領域に対応した複数の CNN を用意する. そして, これら複数の CNN を統合するように上位階層に CNN を用意する. したがって, 未知入力信号パタンの部分領域を下位 CNN により分類し, その結果を上位 CNN によって再構成することで, 最終的な推論結果を得ることが可能となるものである.

(2) 提案する分割階層型 CNN のように学習機械を多数組み合わせることにより, 精度の高い予測や分類を行う手法は一般に集団学習 (Ensemble Learning) と呼ばれ, 近年注目されている. 集団学習では複雑で大規模な学習モデルを用いるのではなく, 比較的単純な学習モデルと計算量が妥当な学習則を用い, 与える例題の重みや初期値の違いなどによって多様な仮説を選び出し, これらを組み合わせることによって最終的な仮説を構成し, 複雑な学習モデルを学習することと同等なことを行うものである.

(3) ところが, 集団学習では多数の学習モデルを構築する必要があり, 個々のモデルに対して適切な学習パラメータ値を割り当てなければならない. 学習パラメータ値の割り

当てには多数の組み合わせが存在する. したがって, 各学習モデルに対して, 適切な学習パラメータ値を考慮しながら割り当てるためには, 試行錯誤的に学習パラメータ値を決定しなければならない. そこで本研究課題では, 学習パラメータ値を適応的に決定するために, 進化的構造適応学習法を適用することを試みた.

(4) 進化的構造適応学習法とは, ニューラルネットワーク (以下, NN とする) の初期パラメータ値 (初期結合荷重やネットワーク構造, 学習係数, シグモイド関数の傾き, 慣性係数) を遺伝的アルゴリズム (以下, GA とする) により決定する手法である. 本研究課題では, 分割階層型 CNN における多数の学習モデルを進化的構造適応学習法によって決定する手法について研究を行う.

3. 研究の方法

(1) 進化的構造適応学習法の改良, その有効性の再検証を行い, 工学的な位置づけを明確にする. 進化的構造適応学習法には, ラマルク型とダーウィン型の2種類の遺伝機構がある. これらの遺伝機構を図1, 2に示す.

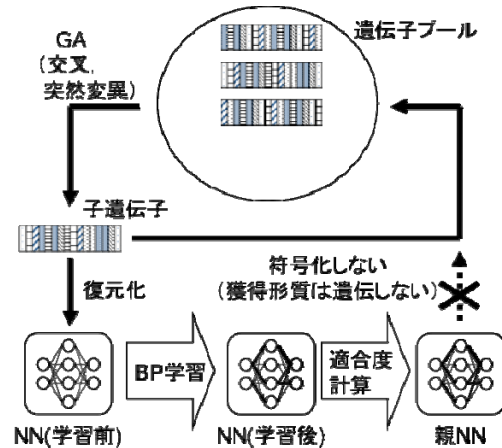


図1. ラマルク型

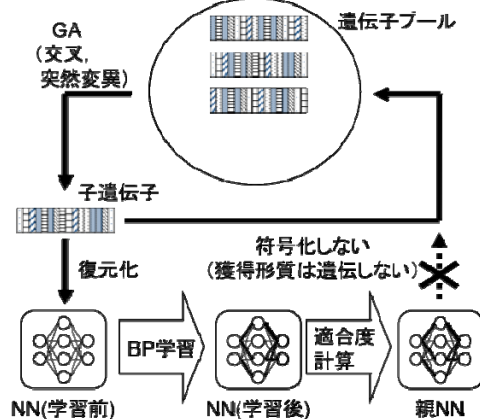


図2. ダーウィン型

著者らの研究では、静的環境ではラマルク型、動的環境ではダーウィン型が適応性に優れているという研究成果を得ている。しかし、変化の大きい、逆に変化の小さい動的環境も存在し、環境の変動に応じてラマルク型とダーウィン型を切り替えながら適応学習させることも有効だと考えられる。そこで、環境変動の度合いと遺伝機構の違いから、環境への適応性を調査する。そして、遺伝機構を使い分けることで適応的に学習を行う手法を提案する。つまり、動的環境の環境変動速度までも考慮した進化的構造適応学習法を提案する。

(2) 提案手法の有効性の検証のため、顔画像の識別実験を行う。そのための顔画像データの取得を次のように行った。表情を変化させた顔画像を得るために図3に示す自作撮影装置による撮影方法を用いる。撮影被験者は、内側を暗くした箱の中に入ったカメラに向かって自分でリモコンを押して写真を撮影する。このとき、顔面固定器を利用してカメラと顔の位置が一定になるようにする。被験者と箱の中のカメラとの間にはハーフミラーを設置し、被験者は自分の表情を見ながら撮影ができるようにする。防音室を用意し、写真撮影時には被験者1人しかおらず、他には誰もいない条件で撮影を行う。なお、被験者には、(1)非常に強くその情動を感じているときの表情をつくること、(2)その情動を感じていることが、第三者にもわかるようにすること、の2点の指示を与える。これらの情報のもとで、喜び、悲しみ、嫌悪、怒り、恐れ、驚きの6つについて4枚ずつ、無表情1枚の撮影を行う。つまり、1人つき25枚撮影することになる。撮影した画像は図4のように顔部分を切り出し、これを顔画像データとする。

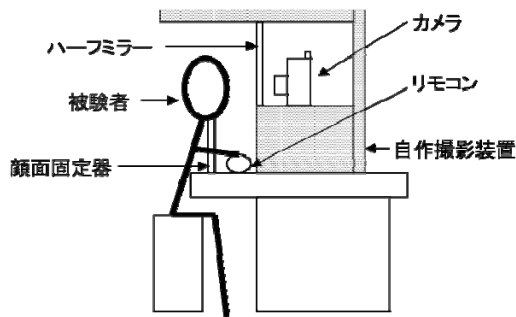


図3. 顔画像取得装置

(3) 以上の研究によって得られた知見をもとに、分割階層型CNNに進化的構造適応学習法を適用する実験を行う。

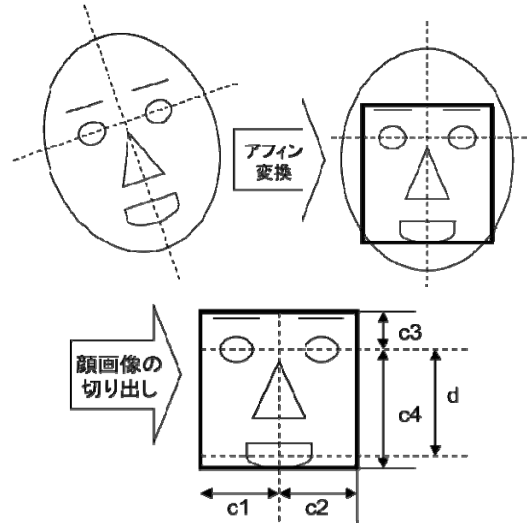


図4. 顔画像の切り出し

4. 研究成果

(1) 進化的構造適応学習法は、GAを用いてニューラルネットワークの最適化を行う手法である。従来、小規模のネットワークモデルへの適用例はあったが、小規模モデルではGAが効果的に機能しているか評価することが困難であった。そこで大規模ニューラルネットワークとして、図5に示す砂時計型ニューラルネットワークに対して進化的構造適応学習法を適用して、顔表情の学習を行った。

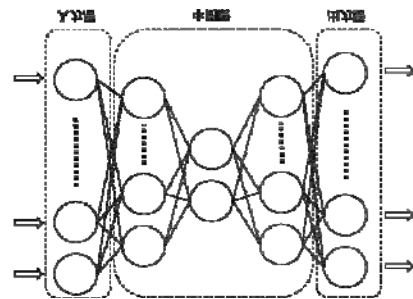


図5. 砂時計型ニューラルネットワーク

顔表情として、喜び、悲しみ、嫌悪、怒り、恐れ、驚きの6つについて4枚ずつと無表情1枚の計25枚の撮影を行った。この撮影を7人に対して行った。撮影した画像は顔部分を切り出し、これらを顔画像データとした。GAが学習に対して効果的に機能しているか調査するため、顔画像の学習に特化した初期結合荷重を探索することを目的として実験を行った。

まず、GAによる砂時計型ニューラルネットワークの最適化を行った。集団内の個体数を20個体とし、1世代において1個体が10,000回のBP学習を行う。初期個体の遺伝子は0を中心とした様々な振れ幅の実数乱数で与える。学習後の平均誤差を基に適合度を算出

し、ルーレット選択、多点交叉、突然変異を施して次世代の子を生成する。次世代へ最優良個体を1個体保存するエリート戦略を用いて、これを20世代まで繰り返した。図7(a)に、各世代における20個体の集団平均誤差の変動を示す。振動しているが、全体として誤差は減少する傾向にある。

次に、GAの代わりにIA(免疫アルゴリズム)を用いて実験を行った。IAでは、各世代における最優良個体を記憶細胞として保存し、これと類似した子が生成されないように抑制することで集団の多様性を維持する。図7(b)に、IAを用いた場合の集団平均誤差の変動を示す。GAと比較して、11世代でいったん誤差が大きくなったが、その後20世代まで概ねIAが上回る結果となった。

GA及びIA、また先行実験で行ったランダムサーチでそれぞれ得られた最優良解を初期結合荷重に用いて、テストデータを学習させた。図8に、学習時の平均誤差の変動を示す。GA及びIAにより得られた初期結合荷重の方が、新しい人物の顔画像を早く学習できていることがわかる。

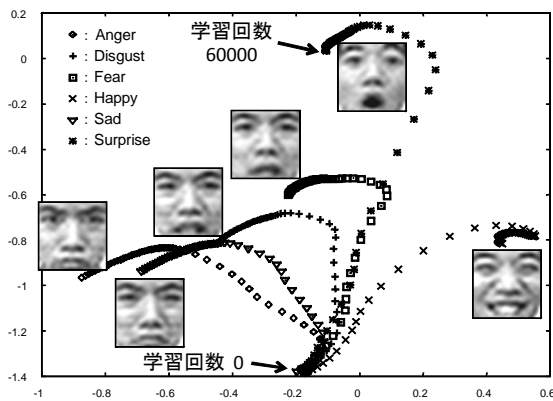


図6. 中間ニューロンの出力の変動

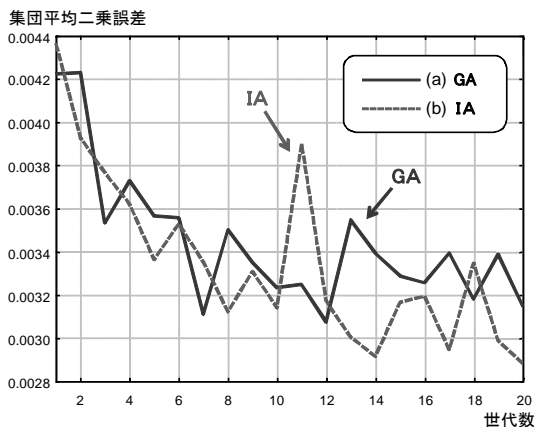


図7. 学習後の集団平均二乗誤差の変動

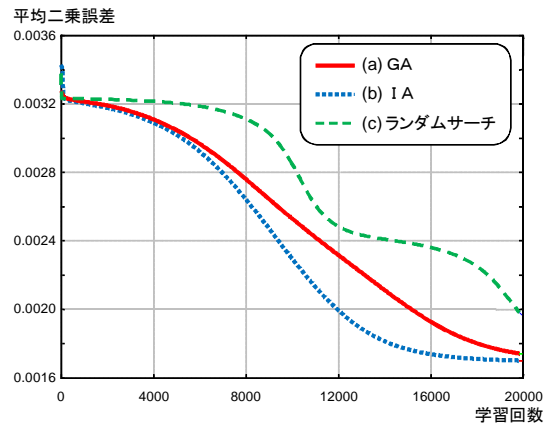


図8. 未知データ学習時の誤差曲線

一般的には、ニューラルネットワークの結合荷重は、ランダムに決定するが、進化的構造適応学習法を用いることにより、ランダムサーチで得たものよりも高速に誤差が収束する初期結合荷重を得ることに成功した。

しかしながら、砂時計型NNの非線形写像によって形成される特徴空間はユーザにとって扱いづらい側面があり、またその正当性についてはほとんど言及されてこなかった。

そこで、砂時計型ニューラルネットワークの中間出力によって形成される特徴空間を調べ、より利便性の高い特徴空間を形成するために、従来とは異なる砂時計型NNの学習法を提案する。具体的には、バックプロパゲーションとネットワーク・インバージョンを組み合わせた、学習中に特徴空間の方向性を指示する手法を提案する。

(2) ネットワーク・インバージョンを用いる事によって、一般的なニューラルネットワークで解く事ができない逆問題を解く事が可能となる。バックプロパゲーションとネットワーク・インバージョンの組み合わせる前に、ネットワーク・インバージョンの問題点を精査し、その改良を行った。

ニューラルネットワークを用いることによって、様々な分類問題を解くことが可能となった。これは、ある入力から出力を得る順問題を解いていることになるが、実際の応用分野では出力を得るにはどのような入力を加えたらよいかという、逆問題を解きたい場合も多く存在する。そこで、逆問題を解くための手法として、ネットワーク・インバージョンが提案されている。

しかしながら、対象となる問題が非線形である場合、ネットワーク・インバージョンでは最急降下法を用いているため、局所解に陥ってしまうことが予想される。著者らは、非線形問題にネットワーク・インバージョンを適用し、局所解に陥るか否か、また局所解に陥る場合にはどのような挙動を示しているか調査、検討を行った。

対象とする非線形関数とそのグラフを図 9 に示す. この非線形関数では, $y=0.2$ となる x の値は, 0.69, 0.94 の二つとなる. 提案した局所解解決手法では, 逆問題の解が複数個ある場合にも対応するため, 初期入力信号 x の値を何度も変更して, ネットワーク・インバージョンを複数回行い, 解が一定の個数に収束した時点で, それらを逆問題の解とする.

その方法とは, 入力信号 x の修正が終了した時点で順方向処理を行う. 次に, 入力信号 x に対するネットワーク出力 o と対象出力 y との二乗誤差を求める. その二乗誤差が定数 β 以下であれば, 入力信号 x を解候補として保存する. 入力信号 x を変更して同様の手順を繰り返し, 対象出力 y に対する最適な入力信号 x を求める. この手法には β の決定方法が確立していない問題があるが, 試行錯誤的に $\beta=0.4$ としたとき, 図 10 に示すように逆問題を解くことに成功した.

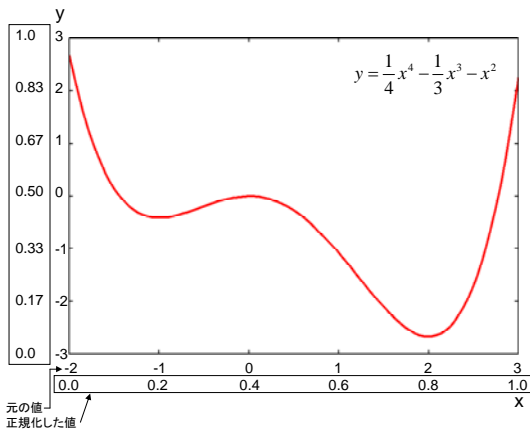


図 9. 対象とする非線形問題

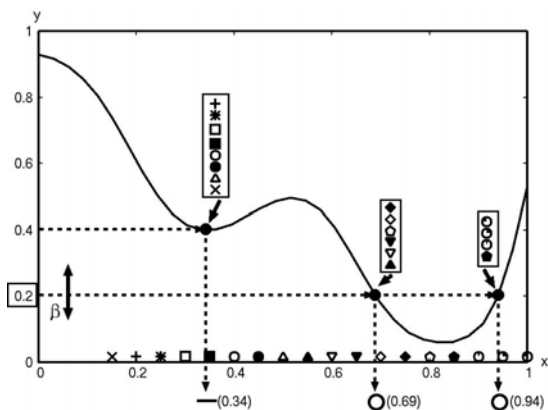


図 10. 出力結果から入力信号を求めた様子

(3) 情報圧縮, 特徴抽出を行うための砂時計型ニューラルネットワークが考案されているが, 砂時計型ニューラルネットワークの非線形写像によって形成される特徴空間はユーザにとって扱いづらい側面があり, またその正当性についてはほとんど言及されてこ

なかった. そこで, 形成される特徴空間を調べ, より利便性の高い特徴空間を形成するために, バックプロパゲーションとネットワーク・インバージョンを組み合わせ, 学習中に特徴空間の方向性を指示する手法を提案する. これにより, 学習中に特徴空間の方向性を指示することにより, データ間の距離関係を多く反映した特徴空間を得ることを目的とする.

バックプロパゲーションとネットワーク・インバージョンを組み合わせた提案手法を DoubleBP と名付け, これを図 11 に示した. これを用いて, 立方体状の座標群を学習させ, 特徴空間を形成する実験を行った. まず, 一般的なバックプロパゲーションのみの学習によって形成された特徴空間を図 12 に示す. また, DoubleBP では恒等写像学習と中間出力修正の BP を各 1 回と考え, 10000 回の BP 学習を行う. 学習完了時に形成された特徴空間を図 13 に示す. これを見ると, 頂点 1 と頂点 7 の組以外は, ユークリッド的な距離関係を 2 次元上にうまく実現していることがわかる. 見方によれば, 特徴空間の歪みを一箇所に集中させた形と捉えることもできる. 図 12 と比較して, 従来手法とは明確に傾向の異なる特徴空間を得ることに成功したと言える.

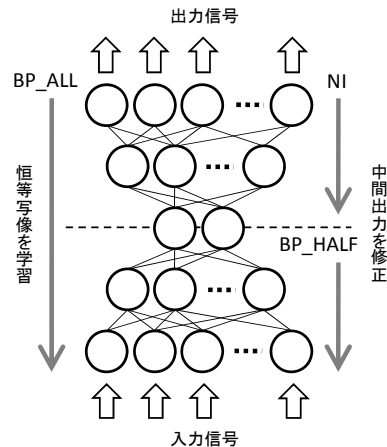


図 11. DoubleBP の概念図

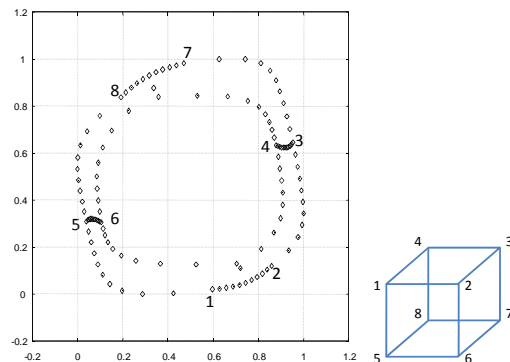


図 12. BP のみにより形成された特徴空間

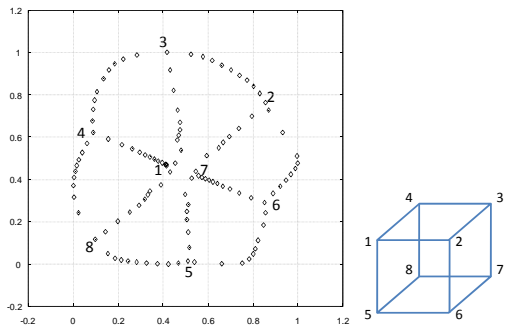


図 13. DoubleBP により形成された特徴空間

(4) Cellular Neural Network の問題点として、識別対象となる課題サイズが大規模になった場合、CNN を構成するセル数も多くなり、漸近安定平衡点までの収束には多大な計算量を必要とするとともに、識別能力も低減することが考えられる。そこで著者らは分割階層型 CNN を提案した。分割階層型 CNN を図 14 に示す。このとき CNN に対して適切な学習パラメータ値を試行錯誤的に割り当てなければならないが、このパラメータ設定の自動化表現が、提案手法一般化の鍵となる。

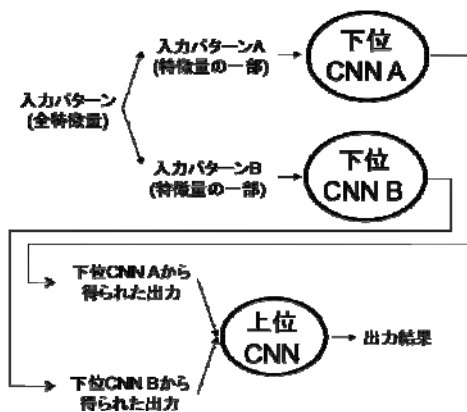


図 14. 分割階層型 CNN

そこで、進化的構造適応学習法を用いて、入力パターンの分割数、学習パラメータを求めた。また、CNN では実数値を扱うことができないため、2 値化する必要がある。しかしながら、255 の実数値を 2 値に量子化すると、多数のセルが必要となってしまう。そこで、砂時計型ニューラルネットワークを用いて、特徴抽出を行った。本研究では、第 3 層目のニューロン数を 10 個として実験を行った。これにより、顔画像データが 10 次元に圧縮できることになる。さらに 10 次元の実数値データを 11 段階に量子化した。11×10 のマトリクス状に配置した CNN により、分類を行った。分割階層型 CNN では、5×11 の下位 CNN を 2 つ用意した。比較対象として、k-means 法との比較を行った。k-means 法での初期クラスター数は 6 と設定した。結果を表 1 に示す。

表 1: 分類実験結果

分類手法	正答率 (%)	時間 (s)
k-means 法	75.6	0.721
CNN	76.1	603.1
分割階層型 CNN	76.1	917.2

k-means 法では、クラスターとデータとの距離計算を行うだけであるため、高速に結果を得ることができた。一方 CNN では若干ではあるが正答率の向上が見られた。しかし長い想起時間を必要とした。分割階層型では、用意する CNN の数が増加するため、全体の計算時間が長くなった。

(5) 本研究によって、以下の成果を得た。

- ・ 進化的構造適応学習法の有効性の検証。
- ・ GA および IA による進化的構造適応学習法による砂時計型ニューラルネットワークの最適化
- ・ 砂時計型ニューラルネットワークに適した初期結合荷重の探索。
- ・ ネットワーク・インバージョンの非線形問題の一部を解決。
- ・ 砂時計型ニューラルネットワークの特徴空間の形成のために DoubleBP の開発。
- ・ 分割階層型 CNN による顔画像の分類。

一方、課題としては、対象問題によっては、分割階層型 CNN の性能を発揮できない場合があること、進化的構造適応学習法が有効に分割階層型 CNN に作用しているか調査する事があげられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 1 件)

1. 川崎直輝, 大枝真一, ネットワーク・インバージョンにおける局所解問題解決手法の提案, 日本高専学会誌, Vol. 13, No. 2, pp. 49-54, 2008, 査読有。

[学会発表] (計 6 件)

1. 小山達也, 大枝真一, 進化的構造適応学習法による砂時計型ニューラルネットワークの最適化, 日本高専学会第 13 回年会講演会, (20070901), 木更津工業高等専門学校。
2. 小山達也, 大枝真一, GA 及び IA による砂時計型ニューラルネットワークの最適化, 情報処理学会第 70 回全国大会, (20080315), 筑波大学。
3. 浦野昌平, 大枝真一, 教師信号あり砂時計型ニューラルネットワークの提案, 電子

- 情報通信学会東京支部学生会研究発表会,
(20090307), 東京電機大学神田キャンパス.
4. 小山達也, 大枝真一, 砂時計型ニューラルネットワークを用いた特徴空間の構築, 情報処理学会第71回全国大会, (20090312), 立命館大学びわこ・くさつキャンパス.
 5. 青木秀人, 大枝真一, 砂時計型ニューラルネットワークによる顔画像の分類, 情報処理学会第72回全国大会, (20100309), 東京大学.
 6. 浦野昌平, 大枝真一, GAを用いた砂時計型ニューラルネットワークの最適化, 情報処理学会第72回全国大会, (20100310), 東京大学.

6. 研究組織

(1) 研究代表者

大枝 真一 (OEDA SHINICHI)

木更津工業高等専門学校・情報工学科・講師

研究者番号: 80390417