

科学研究費助成事業 研究成果報告書

令和 2 年 7 月 7 日現在

機関番号：32692

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00251

研究課題名(和文) 脳の視覚情報処理機構を考慮した深層畳み込みニューラルネットワークの研究

研究課題名(英文) A Study on Deep Convolutional Neural Network Considering the Mechanism of Visual Information Processing in the Brain

研究代表者

菊池 眞之 (KIKUCHI, Masayuki)

東京工科大学・コンピュータサイエンス学部・講師

研究者番号：20291437

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：社会に技術革新をもたらしているCNN(深層畳み込み神経回路)を更に進展させるべく、本研究課題では脳情報処理の観点からのCNNへの機能付加やCNNの性質調査等を行った。まず脳内パターン表現仮説のある中心軸表現の有用性を検討すべく、入力を中心軸変換した際のCNNの性能向上を検討した。また、豊かな構造を持つ脳回路の機能に迫るべく進化的計算によりモデル構造の生成を行った。一方で視覚系の大まかな近似と見なせるU-Netによる図地分離や補完等の視覚機能の実現可能性を調べたり、CT画像からの脊椎抽出を行ったりした。そしてCNNの認識精度と反応分布の関係について調べ、効率的モデル設定の足掛かりを得た。

研究成果の学術的意義や社会的意義

昨今、社会の諸領域に変革をもたらしつつある深層学習の中核技術であるCNNは元々、1960年代の一部の脳科学的知見を工学的にモデリングしたものがベースになっている。更なる脳の知見の投入により、深層学習はもっと進化する可能性も考えられる。万能のように扱われる既存の深層学習も現状では狭い問題設定の枠内でのみ有効性が確認されており、人間の脳ほどの汎用性はまだ獲得されていないこと、そして大量のデータがないと高い能力を発揮することができず、少量のデータのみでも内在する性質を適切に学び取れるヒトに及ばないことなどの難点があり、本研究課題の観点はこれらを解決する糸口を提供し得るものと言える。

研究成果の概要(英文)：In order to refine the CNN (Deep Convolutional Neural Network) causing the innovation on our society, this research project aimed to add novel functions to the CNN and investigated the properties of the CNN from the viewpoint of brain information processing. First, the usefulness of the medial axis representation which is regarded to be encoded in the brain, is examined by applying medial axis transformation to input patterns. In addition, the model structure was generated automatically by an evolutionary computation to approach the function of the brain having rich structures. On the other hand, we investigated the possibility of U-Net approximating the visual system roughly, to simulate visual functions such as figure-ground separation and visual completions, and performed spine extraction from CT images. Moreover, we investigated the relationship between the accuracy of recognition by CNN and the response distribution, and gained a foothold for efficient model setting.

研究分野：脳の情報処理

キーワード：深層学習 畳み込みニューラルネットワーク 脳 視覚系 モデル

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

自然画像認識のコンペティションにて 2012 年に CNN(深層畳み込みニューラルネットワーク)のアプローチが他の方式に対し大差をつけて圧勝したことで注目を集めて以来、多方面で CNN に立脚する AI 技術の活用が進み、各分野に革新的な変化をもたらしており、人間と機器との関係や、人間の行うべきことなどの再定義が迫られるまでも至っている。社会的に多大な功績をあげている CNN であるが、基本的なアーキテクチャーは 1979 年に福島により提案された Neocognitron と大きく変わることはなく、復習種類のフィルタを用いた局所的な畳み込みによる特徴抽出と空間的な情報集約であるプーリングの多段階配置が基礎となっている。これらのモデルが立脚する知見は 1960 年代における Hubel と Wiesel が動物を使った実験を通して見出した、入力刺激に対する位置依存性の高い反応を示す線分検出器である単純型細胞(simple cell)ならびにローカルに位置インバリエントな反応を示す複雑型細胞(complex cell)に基づいており、これらのエッセンスを巧みに工学的観点よりニューラルネットワークの機構として実装したものであることが共通している。その形態を基本とし、約 40 年が経過しているが、畳み込みとプーリングの組み合わせ方のバリエーションは様々な派生形が編み出されてきているものの、質的に顕著な変化はない。従来到達し得なかった領域に昨今踏み込めるようになったのは主として計算機資源の飛躍的進展に伴う量的な進歩が原動力であったと言える。このような CNN も現行のものは万能とは言えない。ごく狭い範囲に限定された問題設定下では人間をも凌駕する認識性能を発揮することもあるものの、人間ほどの広範なタスクへ適応できる汎化能力はまだ獲得されていないこと、そして大量のデータがないとパフォーマンスが上がらず、少量のサンプルのみから法則性を学び取れる人間自身とは大きく異なること、このことに関連し、高性能な GPU を搭載した潤沢なスペックの計算機資源が利用できない場合には学習のために著しく長い時間が掛かってしまうこと、さらには勾配計算を利用した学習の性質の隙を突き、人間には普通に読み取れる画像であっても、特集なノイズを画像に重畳することで CNN に著しい誤認識を生じさせられること、そしてニューラルネットワーク全般に共通することとして、学習により獲得された知識の可読性がゼロに近く、ネットワークの出力の「根拠」を汲み取ることが極めて難しいこと、などさまざまな問題点が残存している。約 40 年間の長きに及ぶ質的進歩の足止め状態から転換し、一步を踏み出すことで、これらの問題のうちの一部でも解決の糸口が見いだせる可能性も考えられる。特に、脳に学ぶことに端を発する CNN に、さらなる脳の知見を投入することで、よりヒトの知性に近づけられるような質的向上が期待される。

2. 研究の目的

本研究課題では脳の情報処理の特質、特に視覚系で見出されている知見を既存の CNN に新たに導入することにより性能の向上が図れるかを試すことを目的の主軸に据える。特に、パターン認識における識別率の向上や、セグメンテーションタスクにおける出力の正確さ、同程度の識別率を達成する学習用データの分量の削減などをどの程度達成できるかを調べる。また、上述の能力向上に向けた取り組みの一環として、脳情報処理の観点より既存の CNN の枠組みがどの程度の能力を有するのかの確認も行う。具体的には脳の視覚系が司る視覚認知にまつわる諸機能を既存の CNN に模擬させたときの遂行の度合いを確認したり、医用画像処理など実用タスクでの達成度を調べたりする。そして、CNN のハイパーパラメータやネットワーク構成を簡便に設定するための方法も模索する。

3. 研究の方法

従来の CNN では 1 つの層内の 1 つの特徴マップ内では細胞同士の間には空間的な近接性(距離)の関係が埋め込まれていたが、層内の特徴マップ同士の間には明示的な関係性は定義されていなかった。一方で現実の脳の視覚系では初期視覚から高次視覚に至るまでトポグラフィックマップの構造が観察されており、空間的に近い位置同士の細胞の最適刺激は類似し、大脳皮質上の位置の連続変化に対し最適刺激も連続的に変化することや、別の方向への位置変化に対しては不連続に変化することなどが知られている。これをモデルの立場で眺めると連続変化は特徴マップ同士に距離が定義され、互いに近い特徴マップ同士は近い特徴を抽出し、遠い位置同士の特徴マップは抽出特徴の類似度が低い、というような構図を想定し、強引に 2D の大脳皮質に埋め込んだ事での連続・不連続な構造になったものと解釈できる。ここではそのような解釈のもとで特徴マップを 2D 座標系に配置し、Kohonen の SOM 同様の近傍関数を定義して近接する特徴マップ同士かつ対応する位置にある細胞同士に hard wired な結合を設置する。このような機構を全層に持たせたネットワークの識別率などのパフォーマンスを測定することを目指した。

次に CNN のアーキテクチャーやハイパーパラメータの設定に関して述べる。初期のモデルでは畳み込みとプーリングをペアにしたものを多段階に配置し、最後に全結合ネットワークを据える形式が一般的であったが、その後アーキテクチャーを自由に策定したモデルが増えてきている。そして、1 つのアーキテクチャーに対してもネットワークのハイパーパラメータをどう設定するかという問題も伴う。こういった問題に対し、本研究課題では進化的計算の枠組みを適用し、パフォーマンスの良いアーキテクチャーやパラメータを模索させることを考えた。さらにはパフォーマンスの高いネットワークのみならず、高いものから低いものまでを生成させたうえで、パフォーマンスレンジごとの反応分布上の性質を吟味することで、大量のパターンを学習さ

せずともネットワークの良しあしを判断できる方法を確立することを目指した。

パターン認識における中心軸(スケルトン)表現の導入については、多数のシルエット状(塗り潰し)オブジェクトパターンを用意した上で、それらに対し、拡大・縮小・回転・剪断・ノイズ付加など様々なタイプの変換それぞれにつき多段階を施したパターンセットを用意した。それら各パターンに対し、外縁輪郭、輪郭内部スケルトン、輪郭外部スケルトンなどを生成し、それらの組み合わせパターンを生成しておく。CNNの学習用には変形等のない基本オブジェクトを用い、識別テスト時には各種変形を適用したものをを用いた。このような設定のもと、パターン種別として、外縁輪郭やスケルトンをどのように組み合わせたものが最も変形等への許容度が高いのか、そしてそれはどの程度かを調査した。

従来のCNNではニューロンを整数グリッド座標に割り当てていたが、実際のニューロンは連続空間中に分布する。グリッド座標採用の場合にはダウンサンプリング時に同一形状の入力に対しても位置に応じて後層側にて見える特徴が変化してしまうという問題があるが、実数座標で緩和されることが期待できる。この方法を導入した際の、主として識別率の観点でのメリットを吟味することを考えた。

そして、従来は一方向性のカスケード状のアーキテクチャーを中心としつつ、LSTMなどのリカレント構造を導入したタイプも考案されたCNNであるが、一方で脳の視覚系では多数のモジュールから構成され、それらモジュール間には双方向に結合が備わっていると同時に、同一視野内でも相互に信号の授受を行うようになっており、リカレント処理の度合いが極めて高いものになっている。このような構造をCNNに導入したときのパフォーマンスについて調べた。

以上の新機軸の導入のみならず、既存のCNNが脳科学的観点よりどの程度の能力を有するものであるかの調査も行った。具体的には、一部が遮蔽されたパターンの輪郭の補完や、図と地の分離、対称性の検出などの視覚機能をシミュレートさせた際のパフォーマンスや、医用画像処理におけるCT画像からの特定組織のセグメンテーションなどでのパフォーマンスを、アーキテクチャ的に視覚系の大まかな近似とみなし得るU-Netを用いて調べた。

4. 研究成果

まず脳内パターン表現仮説のある中心軸表現の有用性を検討すべく、入力を中心軸変換した際のCNNの性能向上を検討した。識別器としてはAlexNetを用いた。結果の一部を図1に示す。入力のタイプの種別ごとに、全変形パターンに共通するような明確な差はないものの、概して物体外縁輪郭やそれに物体外部に形成される中心軸を付加したパターンの識別率が高いケースが多かった。Neocognitronを識別器に用いた場合も類似した状況が観察されたが、そちらでは変形タイプごとのパラメータ最適化を行っていたので、共通のパラメータでの変形追従度を調べる必要もある。

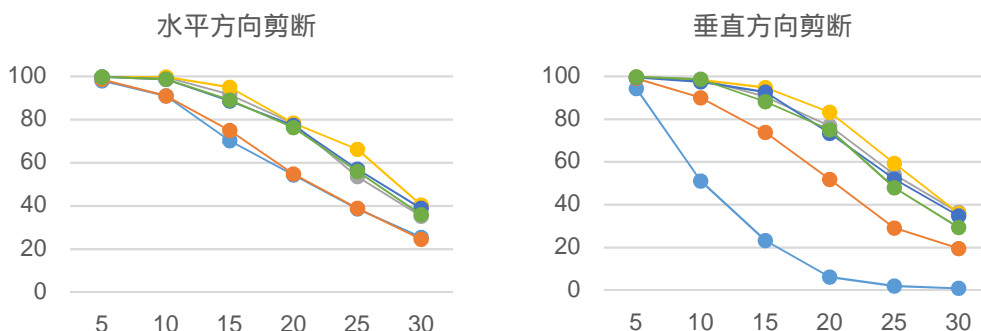


図1 物体外縁輪郭と中心軸表現等の変形に対するパターン認識の許容度の比較(一部を掲載)。左：水平方向剪断，右：垂直方向剪断。横軸：パターン変形度，縦軸：識別率。灰色：物体外縁輪郭，水色：物体領域内部中心軸，黄色：物体領域外部中心軸，オレンジ色：内部+外部中心軸，青色：外縁輪郭+内部中心軸，緑色：外縁輪郭+内部+外部中心軸。

進化的計算によりモデル構造の生成を行わせる取り組みでは、タスクとして、脳の視覚認知において重要な図地知覚の機能のニューラルネットワークの生成を設定し、これまで明らかになっている生理学的知見を断片的制約条件として与えつつ、進化的計算の手法を援用して全体としてコンシステントなモデルとして獲得させることを目論んだ。結果として離合集散する豊かな経路構造を持つネットワークを生成するに至り、パフォーマンスも良好であった。適用したタスクを図地分離とは別のものに置き換えることも可能で、脳機能の理解のためのみならず、汎用的な深層畳み込みニューラルネットワークのアーキテクチャー生成法として活用することも期待される。

一方で、CNNのパラメータを適切に調整する手法を見出すための識別率優良/不良ネットワークの反応様式比較の研究では、学習後の層ごとの特徴抽出結合の類似度を測り、冗長さを吟味した。その結果、性能に応じた冗長さの分布を見出すことができた。これを踏まえることで、多量

の学習用データを全て用いた学習を完遂させずとも、早い段階でネットワークの設定の良しあしを判断できるようにするための足掛かりが得られた。

一方で視覚系の大まかな近似と見なせる U-Net による物体領域のセグメンテーションやアモータル補完、入力パターン中の対称性の検出等の視覚機能の模擬や、CT 画像からの脊椎抽出のタスクを遂行させた。視覚機能の模擬についてはいずれも定性的には実現できる見通しが得られたものの、精度は必ずしも高いものではなかった。ただし、扱ったパターン数がやや少な目であったこともあり、更なる検証が必要である。CT 画像中の脊椎の検出においては 2D でも良好な結果が得られており、さらに x, y, z の 3 軸それぞれの方向で見たときの 2D パターンを使った抽出結果を束ねる工夫を採用したところ、さらに高い精度の結果を得ることができた。

CNN 構造へのトポグラフィック構造の導入や、CNN グリッドの実数化、CNN の全層リカレント化については本研究課題期間終了段階では継続中の状態のままとなった。引き続き成果を導くべく取り組んでゆく。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Naofumi Shigeta, Mikoto Kamata, Masayuki Kikuchi	4. 巻 7
2. 論文標題 Effectiveness of Pseudo 3D Feature Learning for Spinal Segmentation by CNN with U-Net Architecture	5. 発行年 2019年
3. 雑誌名 Journal of Image and Graphics	6. 最初と最後の頁 107-111
掲載論文のDOI（デジタルオブジェクト識別子） 10.18178/joig	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計9件（うち招待講演 2件/うち国際学会 3件）

1. 発表者名 Naofumi Shigeta, Mikoto Kamata, Masayuki Kikuchi
2. 発表標題 Effectiveness of Pseudo 3D Feature Learning for Spinal Segmentation by CNN with U-Net Architecture
3. 学会等名 ICFIP2019(International Conference on Frontiers of Image Processing) (国際学会)
4. 発表年 2019年

1. 発表者名 重田尚郁, 鎌田理詩, 菊池眞之
2. 発表標題 U-Netを利用したCT画像からの脊椎領域抽出タスクにおける立体形状学習の有用性
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会
4. 発表年 2019年

1. 発表者名 菊池眞之
2. 発表標題 ヒトの視覚特性と認知、情報処理のメカニズム
3. 学会等名 技術情報協会セミナー「ARにおける視覚情報の認知、処理のメカニズムと現実空間との整合性」（招待講演）
4. 発表年 2018年

1. 発表者名 鎌田理詩, 菊池眞之, 庄野 逸, 林 勲, 福島邦彦
2. 発表標題 U-NetによるCT画像における脊椎の自動検出
3. 学会等名 電子情報通信学会ニューロコンピューティング研究会技術報告
4. 発表年 2018年

1. 発表者名 Yoshiki Goto, Masayuki Kikuchi
2. 発表標題 Identification and region extraction of symmetry by deep learning
3. 学会等名 NCSP2018 (Nonlinear Circuits, Communications and Signal Processing) (国際学会)
4. 発表年 2018年

1. 発表者名 Mikoto Kamata, Kunihiko Fukushima, Hayaru Shouno, Isao Hayashi, Masayuki Kikuchi
2. 発表標題 Automatic detection of spine in CT image by U-Net
3. 学会等名 NCSP2018 (Nonlinear Circuits, Communications and Signal Processing) (国際学会)
4. 発表年 2018年

1. 発表者名 菅野路哉, 菊池眞之
2. 発表標題 遺伝的プログラミングによる図地分離モデルの自動生成
3. 学会等名 日本神経回路学会第29回全国大会
4. 発表年 2019年

1. 発表者名 菊池眞之
2. 発表標題 凹凸特徴に基づく物体認知の性質とメカニズム
3. 学会等名 第23回日本知能情報ファジィ学会 しなやかな行動の脳工学研究部会 研究会・第121回 日本知能情報ファジィ学会 関西支部例会（招待講演）
4. 発表年 2019年

1. 発表者名 菊池眞之，島森和輝
2. 発表標題 ロバストなパターン認識に対する中心軸表現の効果の検討
3. 学会等名 第36回ファジィシステムシンポジウム
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考