

遺伝的プログラミングによる深層学習モデル構成の一手法

平野 廣美†

† 楽天株式会社楽天技術研究所

1 はじめに

近年、ニューラルネットワークの深層学習 (Deep Learning:DL) において、層の構成を深くしても勾配の喪失を起すことなく学習を達成できるニューラルネットワークの構成 (モデル) が提案されている。特に、Residual Deep Network においては 1k を超す深さが達成されている。この種のネットワークでは、Convolution 層を含む部分構成が繰り返されていて、深い階層が出来上がっている。それらのモデルの構造的な差は主として Convolution 層を含む部分に表れる。しかし、その構成についての指針は特に存在しない。本論では、遺伝的プログラミング (GP) を用いて、適用する問題の特性を捉えた部分構成を進化させ、最終的に DL のモデルを自動構成する手法を提案する。

2 関連研究

ニューラルネットワークの構造を進化させるというアイデアは Neuroevolution において、ニューラルネットワークの学習に遺伝的アルゴリズム (GA) が用いられている [2]。その手法はネットワークの直接コード化と間接コード化の2つに分けられる。前者はネットワークの荷重を GA の遺伝子にコード化し、後者はネットワークの構成ルールをコード化するものである。GP は直接的なコード化には不向きであるが構成ルールを表現することができる。この性質を用いて [4] では、ニューロのノードを2次元のマトリックス状に配置して各ノード間の荷重をノードの座標の関数として与え、与えられた問題をこの関数の同定問題とみなし、関数同定に GP を応用している。

3 提案方式

これまで提案されてきた深層構造のニューラルネットワークモデルは、浅い層構造の繰り返しとなる場合が多い [3]。また、その部分構成中、モデルの独自性が顕著に現れる部分では、層をノードとする信号の分岐と合流を成している。言い換えると単純な重なりではない。モデルの独自性はモデル提案者の経験的アイ

アであるところが大きく、優れたニューラルネットワークとするための指針は存在していない。本論では、この部分構成のノードとなる層を関数として見なし、信号の伝播をそれらの関数を呼び出すプログラムの実行とみなす。このように見ることで、遺伝的プログラミングを適用して、優秀なネットワークを進化的に発見することが可能になる。

3.1 本手法により合成したサブレイヤ構造

既存の ResNet や VGG, GoogleNet モデルは、図1の破線矩形で囲った部分に相当する層構造を挟むようにして更に層が重ねられている。また、この構造が複数組重なって深層構造が形成される。図1の破線矩形で囲まれた部分が本論の手法で合成した層構造の例であり、ResNet タイプと分枝タイプを示している。これら以外に図2のストレートタイプも合成している。本論では、最終的に深層構造として構成するのではなく、中心的な層構造 (ネットワーク構造) を GP により合成して、その前後から定形とみなせる入力層と Full Connection 層で挟む構造とした。

3.2 GP による構造の進化

GP 適用のため、非終端記号と終端記号のセットを準備する。今回の目的では、終端記号として1層の入力層とし、非終端記号としては深層学習フレームワーク (Keras/Theano) が提供する層 (レイヤ) から以下の4種類とした。GP ではこれらの記号をノードとした木構造を合成する。木の初期サイズは、ノード数を3~12の範囲で中央値が最大となる分布から決定した。

- Terminal = {InImage},
- Non Terminal = {Merge, Convolution2D, Max-Pooling2D, AveragePooling2D}

4 実験と考察

実験には、Cifar-10 (Training : 50,000, Validation : 10,000) を用いて合成した層構造について能力を調べた。DL のフレームワークとして Keras/Theano を、GP フレームワークとして DEEP を用いた。実行時のパラメータを、表1と表2にまとめる。

A Model Construction Method Using the Genetic Programming in the Deep Learning

†Hiromi HIRANO

†Rakuten Institute of Technology, Rakuten, Inc.

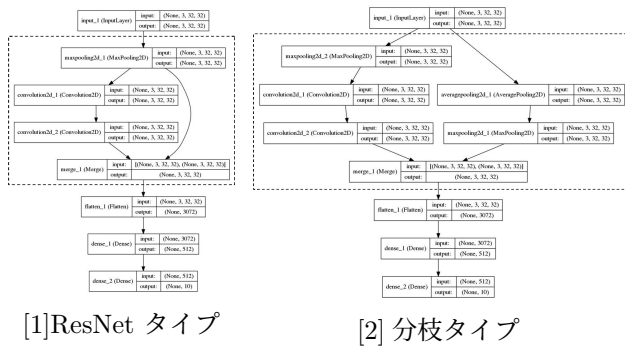


図 1: 画像の変換例

表 1: GP parameters

DEEP	
世代数	15
個体数	20
交叉確率	0.9
突然変異確率	0.1

表 2: DL parameters

Keras/Theano	
Epochs	200
Mini batch	32

4.1 実験結果

表 1 のパラメータ条件下で GP により合成した固有の層構造を持つ 20 個体を生成し、それら個体は表 2 のパラメータで Keras により Cifar-10 の学習を行う。図 2 は同学習で得られた最良な個体の層構造を示し、矩形内の 6 層が本手法で合成した部分となる。また図 3 は、同最良個体の学習曲線を 200 Epochs の間プロットしたグラフである。図 3 中の 4 本の曲線は、上から、学習正解率、テスト正解率、テストロス、学習ロスを示す。図 4 は 7 世代間の世代別平均正解率の遷移をプロットしたものであり、GP による進化を示している。実験における進化と個体の学習の結果、最良個体のテスト正解率は、200 Epochs 後、0.724 となった。この結果は高い値ではないが、より層を深くすることで改善できると考える。

5 おわりに

本論では、GP により進化的に合成した層構造を含む階層型のニューラルネットワークの能力を調べた。深層構造のニューラルネットワークは、1K を超す深さのものが提案されているが、本論では 10 層程度でしか検証していない。しかしながら、GP により個体の持つ層構造を仮説として多数提示し、個体の学習結果を評価することで、優秀な個体(層構造)へと進化させることが確認できた。今後は、(1) 合成する木のサイズを増やすことで多様性が大きい層構造を生成すること、(2)

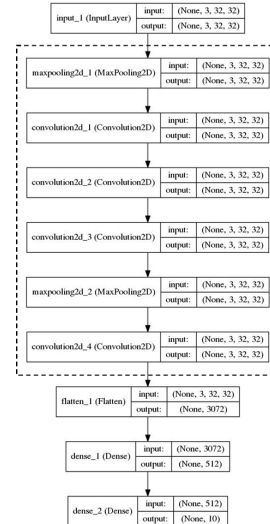


図 2: 最良個体の層構造

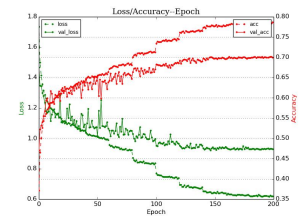


図 3: 最良個体の学習曲線

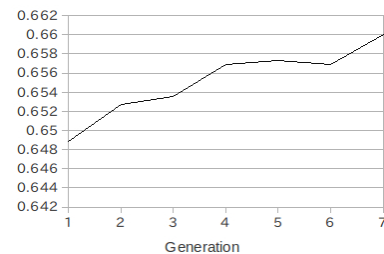


図 4: 平均正解率の進化遷移

更に深い層 (100 層以上) 構造への拡張をおこなうことで、適用する問題に対応した学習能力の高い深層構造の発見を自動化していく予定である。

参考文献

- [1] Sreenivas Sremath Tirumala "Implementation of Evolutionary Algorithms for Deep Architectures" *AIC*, 164-171, 2014
- [2] Omid E. David, Iddo Greental "Genetic Algorithms for Evolving Deep Neural Networks" *GECCO'14*, 1451-1452, 2014
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun "Deep Residual Learning for Image Recognition" *CVPR 2016*, 770-778, 2016
- [4] 海内映吾, 伊庭斉志 "遺伝的プログラミングを用いたニューロ進化に関する研究" 東京大学工学部, 2016