

## 勾配降下法を用いたニューラルネットワークにおける局所受容野の最適化手法の提案

Proposal of optimization method of local receptive field using gradient descent

田爪 聡<sup>1</sup> 大知 正直<sup>1</sup> 坂田 一郎<sup>1</sup> 森 純一郎<sup>1</sup>

So Tazume<sup>1</sup>, Masanao Ochi<sup>1</sup>, Ichiro Sakata<sup>1</sup>, and Junichiro Mori<sup>1</sup>

<sup>1</sup> 東京大学

<sup>1</sup>The University of Tokyo

**Abstract:** In this research, we try to design a method to search the optimum model according to individual data. In the existing structure search algorithm, the applicable types of data are limited. In this study, we will input the receptive field of the convolutional layer. It is possible to apply various types of data to the convolutional neural network. In concrete terms, by interpreting the receptive field as an index that associates inputs and weights in the convolutional layer. We designed a new layer expressing the receptive field in a matrix and also made it possible to learn receptive field by gradient by relaxing its matrix to continuous value. The result of the experiment. The proposed method for the data whose data structure is unknown. By using the method proposed in this paper, it is expected that the spread of IoT and the sensor. Relative big data that is expected to increase in the improvement of technology, the neural network is expected to be effectively applied.

### 1 はじめに

近年ニューラルネットワークの研究が盛んに行われており、産業分野への導入事例も年々増加している。しかし研究分野で扱われるデータと比べて、実応用で用いられるデータの多くはデータの構造に関する知見が十分に得られていないケースが多い。それらの理由からモデルの設計にコストがかかり、導入が進んでいないという課題がある。そこで本研究では個別のデータに応じて最適なモデルを探索する手法の設計を試みる。

個別のデータに対応する最適なモデルを設計する手法としてメタラーニングに関する研究が盛んに行われている。それらの研究の多くはレイヤー間の接続を最適化する手法を用いており、画像処理や自然言語処理など一部のデータに対して優れた性能を示している。しかし既存の構造探索アルゴリズムでは適用可能なデータの種類の数が限られている。そこで本研究では畳み込み層の受容野を入力に対して適切に対応付けることで、様々な種類のデータを畳み込みニューラルネットワークに適用することを可能とした。具体的には畳み込み層において入力と重みを対応付けるインデックスとして受容野を解釈することで受容野を行列で表現する新たな層を設計した。またその行列を連続値に緩和することで勾配による受容野の学習を可能とした。

MNIST を用いた精度評価実験ではこれまで画像分野で用いられてきた畳み込み層の受容野と一部一致する受容野が学習によって得られ、精度評価に関しても同

等の性能を観測した。また気象庁のホームページから収集した日本の過去の気象データを用いた精度比較実験では小さいネットワークで既存の分類手法より優れた結果を示した。

これらの結果からデータ構造が未知のデータに対して提案手法が有効であることを示した。

### 2 関連研究

#### 2.1 畳み込みニューラルネットワークの受容野に関する研究

Yu らは画像データに対して等間隔空けたピクセルを受容野にすることでより広域の情報を畳み込む dilated 畳み込み層を提案した [1]。Jeon らは画像データに対して中心となる受容野以外の受容野の位置を可変とし、位置も含めて学習することで受容野形状を最適化する手法を提案した [2]。

Niepert らはグラフに対して畳み込みニューラルネットワークを適用する手法の提案をした [3]。入力となるノードをグラフデータに関する知見を元に選択し、それらのノードを受容野として畳み込みニューラルネットワークへの適用を行った。

#### 2.2 ニューラルネットワークの構造探索アルゴリズム

ニューラルネットワークの構造探索の手法として様々な手法が挙げられる。Zoph らは様々な構造のモデルを

生成する RNN を生成されたモデルの分類精度を報酬として、強化学習を用いて学習する手法を提案した [4]. また Liu らは離散値であるネットワークのトポロジを連続値に緩和することで勾配降下法を用いたネットワークの構造最適化を行う手法を提案した [5]. ニューラルネットワークのモデルを最適化する手法の多くはレイヤー間の接続を最適化する. それらの手法は画像分野や自然言語処理分野で特に高い性能を示している.

### 2.3 本研究の位置づけ

モデルの構造に関する研究はレイヤーの接続に関する研究と畳み込み層の受容野に関する研究が主に挙げられる. また構造の自動探索に関する研究も多数発表されているが, そのほとんどはレイヤーの接続の最適を扱っている. そこで本研究では, 構造探索手法ではあまり研究が行われていない受容野の最適化に関する研究を行う. また計算量を抑えるために勾配降下法を元に構造の探索を行う.

## 3 勾配降下法を用いた畳み込み受容野の最適化

### 3.1 緩和畳み込み層の提案

本節では受容野テンソルを導入した緩和畳み込み層について説明する. 緩和畳み込み層は受容野行列  $L$  とカーネルの重み  $W$  から構成されており, 入力  $I$  に対して 2 回の行列計算を施す. まず 1 回目の行列計算は入力  $I$  と受容野行列  $L$  のテンソル積を計算している. これは入力の各成分とカーネルの重み  $W$  の各成分を対応付ける操作と対応している. そして得られたテンソルとカーネルの重み  $W$  のテンソル積を出力に用いている. 受容野行列が従来の画像の畳み込み層と同様の受容野を表している場合,  $I \otimes L$  の  $n$  軸方向のベクトルは近隣ピクセルの値を成分にもつベクトルになる. また緩和畳み込み層の順伝播の式を式 1 で示す.  $O$  は出力テンソル,  $I$  は入力テンソル,  $W$  はカーネルの重み,  $L$  は受容野テンソルをそれぞれ表す.

$$O = I \otimes L \otimes W \quad (1)$$

### 3.2 受容野のテンソルの最適化の概要

前節の定義に基づいた場合, 受容野テンソルは離散値となるが離散値では勾配を用いた最適化手法を適用

することができない. そこで本手法では受容野テンソルを学習する段階では  $L$  と同型のテンソル  $R$  に対し被畳み込み軸方向に Softmax 関数を適用することで, 受容野テンソルの値を 0 から 1 の間の連続値として扱った. Softmax 関数はベクトルに対し適用される関数であり入力ベクトルを  $x$ , 出力ベクトルを  $y$  とすると,  $y$  の各成分は以下の式 2 で計算される.

$$y_i = \frac{e^{x_i}}{\sum_{j=1}^N e^{x_j}} \quad (2)$$

このとき  $N$  は  $x$  及び  $y$  の成分数である. 本研究では  $L$  と同型のテンソル  $R$  を直接学習することで受容野テンソルを間接的に学習する. Softmax 関数で計算された受容野テンソルを以降では連続受容野テンソル  $L_c$ , テンソル  $R$  を受容野元テンソルと呼ぶことにする. またネットワークを学習する段階では, 受容野テンソルは被畳み込み軸方向で one-hot である必要があるため, 受容野テンソルを受容野元テンソルを用いて計算する. このとき,  $R$  に対して被畳み込み軸方向に Max 関数を適用する. Max 関数は以下で定義する.

$$y_i = \begin{cases} 1 & (i = \underset{i'}{\operatorname{argmax}} x_{i'}) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

$R$  を元に Max 関数で計算された受容野テンソルを離散受容野テンソル  $L_d$  とする.

図 3.1 に学習の概要を図示する.  $W$  の更新は上の枠線で囲まれた領域で示した.  $b_{train}$  がネットワークに入力されることで  $\mathcal{L}_W$  が計算される.  $W$  は  $\mathcal{L}_W$  を最小化する方向に更新される. また下の枠線で囲まれた領域で  $R$  の更新を示した.  $R$  の更新段階ではネットワークに  $b_{train}$  と  $b_{val}$  を入力することで  $\mathcal{L}_R$  を計算する.  $R$  は  $\mathcal{L}_R$  を最小化する方向に更新される.

$W$  と  $R$  の両方の更新の段階において  $b_{train}$  と  $R$  と  $W$  は共有される.

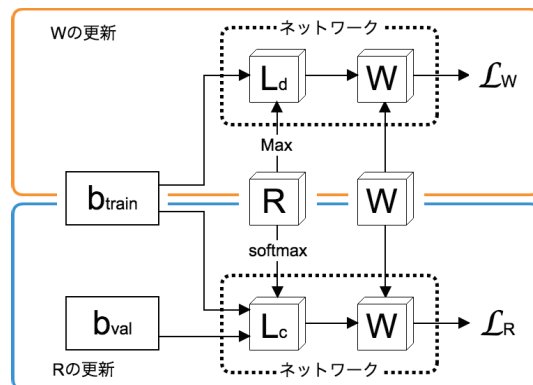


図 3.1 提案手法の学習の概要図

### 3.3 損失関数

次に本手法で用いる損失関数について説明する．本手法ではネットワークのパラメータの学習に用いる損失関数と受容野元行列の学習で用いる損失関数に異なる関数を導入した．パラメータ  $W$  の学習にはソフトマックスクロスエントロピー関数を用いている．受容野元行列  $R$  の学習では Liu らの考案した損失関数を用いた [5]．その関数を次の式 4 で示す．

$$\mathcal{L} = \mathcal{L}_{val}(W - \gamma \nabla_W \mathcal{L}_{train}(W, R), R) \quad (4)$$

この損失関数を最小化するような受容野元テンソルを計算することで，損失関数の更新幅が大きくなるような受容野が得られる．

## 4 実験と結果

### 4.1 実験の詳細

本研究では実験結果を可視化することで受容野を定性的に評価するために緩和畳み込み層の構造を多少変更した．本実験では受容野行列の解釈を考察するという目的を達成するために受容野行列に制約を課した．その制約は受容野行列の受容野軸方向の成分のうち1つを単位行列に固定するという制約である．また本研究では固定された単位行列によって定められた受容野を受容野の中心と呼ぶこととする．

### 4.2 MNIST を用いた実験

本実験では提案手法の妥当性を評価するために，画像分類タスクで頻繁に用いられる MNIST を用いて精度評価の実験を行った．

本実験では受容野元行列を学習し，その行列を定数としてネットワークのパラメータを学習した．図 4.1 に学習で得られた受容野の一部を可視化した．

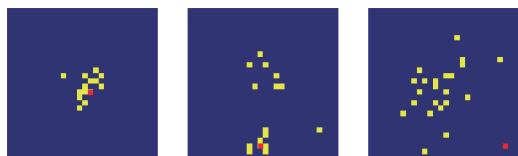


図 4.1 学習で得られた受容野

受容野の中心を定め，それと同時に畳み込まれている他のピクセルを同じ図にまとめた．一番左上のピクセルを (1,1) とした時，それぞれの受容野の中心はそれ

ぞれで (15,15),(13,25),(25,25) である．他の受容野は白色で示してある．次に比較実験の結果を示す．比較実験として，線形結合層のみから作られているネットワークと画像の畳み込みニューラルネットワークも同様に実験を行った．精度は以下の表 4.1 になっている．精度は CNN と近い結果が得られた．

表 4.1 各種手法による MNIST の分類精度

手法	分類精度
提案手法	97.37%
CNN	97.47%
Linear	96.21%

### 4.3 Shuffled MNIST を用いた実験

次に手書き文字画像を並び替えたデータセットに対して提案手法を適用した．画像の畳み込みニューラルネットワークは隣接ピクセルの値を畳み込むことでその精度を向上させたが，その受容野は画像処理の知見に基づき設定されたものであった．知見が十分に得られていないデータや位置情報を持っていないデータに対しては畳み込みの受容野を決定することが難しいという課題がある．そこで Shuffled MNIST では擬似的にそのようなデータを再現した．

各手法の精度を表 4.2 にまとめた．テストデータを用いた精度比較において CNN の精度は著く下がっていることが確認できる．しかし提案手法はデータの加工前と同等の精度を示した．

表 4.2 各種手法による Shuffled MNIST の分類精度

手法	分類精度
提案手法	97.49%
CNN	95.89%
Linear	96.27%

### 4.4 気象庁データベースを用いた実験

本実験では気象庁データベースからダウンロードしたデータに対して提案手法を適用した．気象庁データベースからダウンロードしたデータは日本の各地点の気象状況を格納したデータセットとなっている．そこで本実験では日毎の気象データからその日の月を推定する問題に対して提案手法を適用した．

また各種手法の精度を表 4.3 にまとめた．それぞれの手法のハイパーパラメータはグリッドサーチで選ばれたものを用いた．



表 4.3 各種手法の日本気象データセットの分類精度

手法	分類精度
提案手法	<b>80.08%</b>
Linear	76.62%
線形カーネルを用いた SVM	75.79%
RBF カーネルを用いた SVM	75.86%
決定木	61.61%
ランダムフォレスト	71.19%
k-近傍法	58.54%
ロジスティック回帰	65.36%

## 5 考察

本節では実験によって学習された受容野について考察することで提案手法の妥当性について議論する。

MNIST データセットに対して受容野を学習したところ中央付近のピクセルでは近傍のピクセルを畳み込んでいた。また画像の外側を受容野の中心とする受容野は広がりを持って分布していた。この結果より MNIST に対して受容野の学習を行った結果、画像の中央付近では従来の画像畳み込みと類似の受容野が学習で得られた。また画像の外側付近で受容野が広がりを持っていた理由として、MNIST はクラスの特徴を表すピクセルが中央付近に集めるようにデータセットが作られている。つまり画像の外側では局所的な特徴量を捉える必要がない。そこで外側のピクセルに関しては画像の全体を畳み込むことの方が MNIST というデータセットの構造を捉えられているのではないかと考えられる。

次に日本の気象データに対して提案手法を適用した結果について議論する。受容野の中心が北海道の各地または那覇の時、その他の受容野も同じ地方に集中した。しかしそれ以外の地域を受容野野中心とした時、受容野は南北に幅広く分布していた。気象予報士の方にお話を伺ったところ、日本各地のその日の気象データから月を予測するというタスクにおいて、各月の特徴をもっとも捉えている気象条件の 1 つに前線の位置が挙げられるとのことである。前線とは気団同士の境界上に観測され、気象的に不連続境界となる。畳み込み層は受容野内の構造的特徴を捉える役割を持つことから、受容野を南北に広く分布させることで前線の位置をモデルが獲得しているのではないかと考えられる。そして那覇と北海道は前線の南限と北限の外側にあり、気候区分も日本の他の地域とは異なる。そこで異なる気候区分が受容野に含まれないように受容野の学習が進んだのではないかと考えられる。

## 6 結論

本研究の目的は、知見が十分に研究されていないようなデータに対して畳み込みニューラルネットワークの受容野を緩和、学習することによってそのデータに対して最適な受容野を獲得することであった。この目的に対する本研究の結論について述べる。

MNIST の実験では得られた受容野は画像の中心付近では近隣ピクセルを畳み込んでいた。画像の端付近では受容野が広がりを持っていたが、これは MNIST データセットにおける文字が中央に位置するという特徴を捉えたためだと解釈でき、本手法がデータの特徴を捉えるという点で有用であることを示せた。

気象庁のデータベースから作成したデータセットを用いて実験で学習された受容野を観察したところ地理的、気候的に特徴的な箇所に集中しており、タスクに対する最適なデータの構造を捉えられていた。また実験の結果、既存手法に比べてより高い精度を示しており、学習された受容野が精度向上に寄与したことが示された。

以上から、本研究は特にデータに対する知見が少ない場合において有効性を示し、データの構造を自動で捉えてより効率的に学習するネットワークの構造を探索する手法を提案した。

今後の展望として、受容野元テンソルのパラメータ数の削減を行うことが考えられる。例えば  $k$  次のテンソル分解を適用することでパラメータの数を減らすことができ、より効率的に学習が進むのでは無いかと考えられる。

## 参考文献

- [1] Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.
- [2] Yunho Jeon and Junmo Kim. Active convolution: Learning the shape of convolution for image classification. *CoRR*, Vol. abs/1703.09076, , 2017.
- [3] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. In *International conference on machine learning*, pp. 2014–2023, 2016.
- [4] Barret Zoph and Quoc V Le. Neural architecture search with reinforcement learning. *arXiv preprint arXiv:1611.01578*, 2016.
- [5] Hanxiao Liu, Karen Simonyan, and Yiming Yang. Darts: Differentiable architecture search. *arXiv preprint arXiv:1806.09055*, 2018.