

フィルター学習を行う畳み込み球面自己組織化マップを用いた画像のクラスタリング

福田千恵¹ 堂藺浩²

概要: 近年は情報化が拡大し、情報の整理が難しくなっている。したがって、よりわかりやすく、より有用な情報の自動分類が求められている。そこで、本研究では、フィルター学習を行う畳み込み球面自己組織化マップを用いた画像のクラスタリング法を提案する。自己組織化マップとは分類アルゴリズムの1つであり、データ間の関連性を視覚化できる。データを学習して作成したフィルターを使用した畳み込みニューラルネットワークと自己組織化マップを組み合わせることで、画像データのクラスタリング精度が向上した。本研究では、自然画像セットと MNIST 手書き画像を使用したクラスタリング結果を比較する。

キーワード: 画像特徴抽出, 画像分類, ニューラルネットワーク

Clustering of Images using Spherical Convolution Self-Organizing Map with Learning Filter

CHIE FUKUDA^{†1} HIROSHI DOUZONO^{†2}

Abstract: Recently, it is difficult to organize information because of the huge amount of information. Thus, it is required to organize the information automatically as to be more uncomplicated and more useful. For this purpose, an automatic classification method using Spherical Convolution Self-Organizing Map with Learning Filter is proposed in this research. Self-Organizing Map is one of the classification algorithms and it can visualize relations among data. The convolution with learning filter is introduced to Self-Organizing Map to improve the clustering accuracy of image data. The clustering results using natural image set and MNIST handwriting images are compared in the experiments.

Keywords: Image feature extraction, Image classification, neural network

1. はじめに

近年、スマートフォンやタブレット端末などの普及によりインターネットの利用者が拡大し、音声や文章、画像などさまざまな情報のやり取りが活発に行われるようになった。しかし、情報が多くなるにつれ、その情報を系統立て整理することが困難になってきている。そこで、今回は画像データに注目し、画像の特徴に基づいて分類し、表示するシステムの開発を試みた。

画像認識の手法としては、ニューラルネットワークが広く用いられている。ニューラルネットワークとは、人間の脳を構成する細胞であるニューロン(神経細胞)をモデルにした情報処理システムのことである。ニューラルネットワークには、様々なアーキテクチャが存在するが、本研究で用いたのは、SOM(SOM:Self-Organizing Map)[1]と畳み込みニューラルネットワーク(CNN: Convolutional Neural Network)[2]である。

SOMは、高次元の入力データを自動的に分類し、入力データの類似性や順序性に従って、低次元空間上に写像する

ことができる。写像して得られたマップは、入力データ間の関係が直感的にわかりやすく視覚化されたものとなる。高次元データの可視化、クラスタリング、予測など、さまざまな分野で用いられている。

CNNは、人間の視覚をモデルに考案されており、限定された領域ごとに像を認識するようになっている。特定の形状に対応する畳み込み層と、形状の空間的なずれに対応するプーリング層から成り立つ。画像認識、自動運転、監視カメラ、オンラインショッピングの商品検索など、さまざまな分野で活用されている。

本研究では、CNNで画像データの特徴を抽出し、その特徴を用いてSOMで分類することで、クラスタリングの精度向上を目的とした。

2. 球面 SOM

SOMは入力データ同士の相互関係を、平面や球面などの低次元空間上に表現することができる。平面よりも球面の方が、より柔軟に入力データ同士の関係性を表現できるた

¹ 佐賀大学大学院
Saga University Graduate School

² 佐賀大学
Saga University

め、本研究では球面マップ上に写像できる球面 SOM を用いる。

一般的な SOM の構造は、図 1 に示すように、 m 個のノードを持つ入力層と n 個のノードを持つ競合層から成り立つ。競合層上の i 番目のノードが有する参照ベクトル w_i (結合荷重) は、入力データ x の値に近づくような学習を繰り返すことで、似た入力データがマップ上の近くに配置される仕組みになっている。

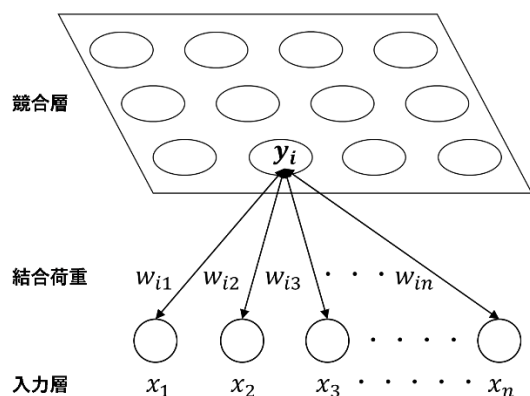


図 1 一般的な SOM の構造

球面 SOM の写像領域は、正多面体の中で頂点の数が最も多い正二十面体を用いることで球面を表現する。正二十面体の面はすべて三角形であることを利用し、それぞれの三角形の辺の中点を新たに多面体の頂点として追加する。そして、近くに配置された頂点をそれぞれ直線で結ぶと再び三角形が得られる。これを繰り返すことにより、三角形の面でできた球に近い多面体を作ることが出来る。また、頂点の数は k 回の分割で $2+10 \times 4^k$ 個となる。作成した球面を図 2 に示す。

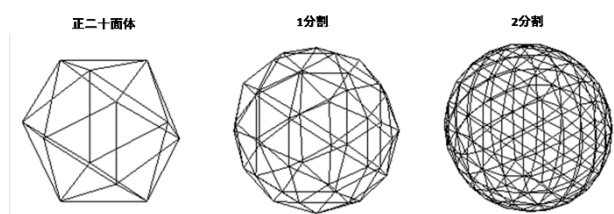


図 2 正二十面体の分割

以下に球面 SOM のアルゴリズムを示す。

- 参照ベクトル w_i の初期値を、乱数を用いて設定する。また、入力層上 j 番目のノードに入力ベクトル x_j を与える。
- 競合層で各ノードの参照ベクトルと入力ベクトルとの距離を計算し、距離が最小のノードを求める。これを勝者ノード N_{k*} と呼ぶ。式を(1),(2)に示す。

$$N_{k*}^j = \min_k f_j(k) \quad (1)$$

$$f_j(k) = \sqrt{\sum_{i=1}^n (x_i - w_{ji})^2} \quad (2)$$

- 勝者ノードとその近傍ノードの参照ベクトルを以下の式に従って更新する。

$$h(P_{jx^*}, P_{jx}) = \exp\left(\frac{-|P_{jx^*} - P_{jx}|}{\theta^2}\right) \quad (3)$$

$$\Delta w_{ji} = \alpha(t) h(P_{jx^*}, P_{jx}) (x_i - w_{ji}) \quad (4)$$

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} \quad (5)$$

ここで、 P_{jx^*} は勝者ノードのマップ上の位置、 P_{jx} はその近傍ノードの位置を示す。 θ は初期近傍範囲の広さを表す立体角、 α は学習率係数を示し、学習回数 t を重ねるごとに小さくなる係数である。 $t: t+1$ として B にもどり、 t が定めた学習回数 T に達するまでこの過程を繰り返す。

3. 球面畳み込み SOM

先行研究[3]では、球面 SOM で画像をクラスタリングする際、固定のフィルター画像を用いた畳み込み処理とプーリング処理を行ったデータを SOM に入力し、学習させている。以下に、画像の処理の手順を示す。

- 図 3 のように、対象となる画像を一部重ねながら 49 分割する。



図 3 画像の分割

- RGB の色ごとに、対象の画像とフィルター画像間の相関係数を計算する。これを畳み込みという。フィルター画像は 27 枚用いた。
- 図 4 のように、分割した画像で隣接する 3×3 の範囲内の画像において、B で求めた相関係数の最大値をとる。これを、対象画像の全領域(25 か所)についてそれぞれ求める。この処理をプーリングという。

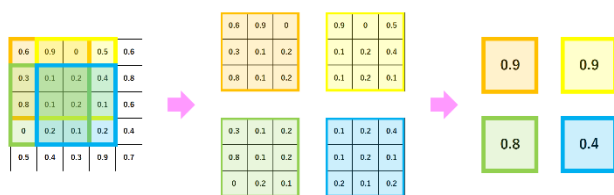


図 4 画像の分割

先行研究の手法により得られた結果を図 5 に示す。入力画像は、「自然」と google で検索して得られた 30 枚の画像である。球面の左手前側には空を含めた画像が、それ以外では大部分を緑が占める画像が配置されていることが分かる。しかし、球面の右下に緑の画像に囲まれた配置で青い風景の画像が配置されていることが分かる。RGB を考慮し、空を含めた画像の近くに配置されると予想していたが、望ましい結果にはならなかった。



図 5 球面畳み込み SOM の実行結果

4. フィルター学習型球面畳み込み SOM

また、先行研究として、フィルターの学習を SOM で行う平面畳み込み SOM[4]の研究が行われている。また、この研究においては MAX プーリング処理において、相関係数のみでなく、SOM で学習されたフィルター画像のマップ上の位置情報も活用や、MAX プーリング処理を SOM で行うことでクラスタリング能力の向上を行っている。また、畳み込み処理の多層化を行い、深層学習も可能となっている。

球面畳み込み SOM の結果が望ましいものにならなかった原因の一つとして、フィルター画像が固定のものであったことが考えられる。そこで、本研究では、フィルター学

習型畳み込み自己組織化マップと同様に、画像を SOM 上で学習させることで得られる特徴画像をフィルター画像として用いる手法を提案する。

SOM で学習して得られる参照ベクトルは、複数の入力データが共通して持つ特徴の影響を大きく受ける。そこで、画像を SOM で学習し、得られた参照ベクトルをフィルター画像として用いることで、より入力画像同士の相互関係に注目した形のクラスタリングを試みる。

フィルター画像の作成に用いる SOM は二十面体上に写像することにした。面は競合層上のノードとなるため、ノード数は 20 個である。SOM には対象画像を 49 分割した画像を入力し、RGB 各色について学習する。得られるフィルター画像は $20 \times 3 = 60$ 枚となる。また、相関係数をただ単に 1 次元のベクトルとして扱うのみでなく、球面上に配置された位置情報も活用し、球面の畳み込みを行う、より高度なクラスタリングを可能にすることを目的とする。

5. WEB 画像を用いた画像分類実験

5.1 実験方法

球面畳み込み SOM とフィルター学習型球面畳み込み SOM の比較を行う。入力画像は google で「自然」と検索して得られた 30 枚と、「世界遺産」と検索して得られた 30 枚を用い、2 パターンのクラスタリングを行う。画像のサイズは 100×100 ピクセルでそれぞれ 49 枚に分割する。分割された画像のサイズは 40×40 ピクセルである。フィルター画像は球面畳み込み SOM は 24 枚、フィルター学習型球面畳み込み SOM は 60 枚である。クラスタリング結果を示す競合層は、1280 のノードを持つ。学習回数は 300 回、学習率係数は 0.8、球の半径は 250、近傍領域は 220 と共通している。

5.2 実験結果と検討

本研究で用いた固定フィルターと SOM で学習して得られたフィルターは図 6 の通りである。SOM で学習したフィルター画像はほとんど均等な色や、グラデーションがかかった色になった。

「自然」と検索した画像を分類した結果、図 7,8 のようになった。先行研究と比較し、フィルター学習型球面畳み込み SOM は、より色の特徴でのまとまりが強くなったクラスタリングがなされているが、どちらもテクスチャーが捉えられていないことがわかる。

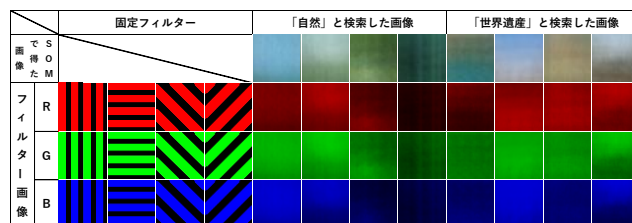


図 6 フィルター画像



図 7 球面畳み込み SOM で自然画像を分類した結果

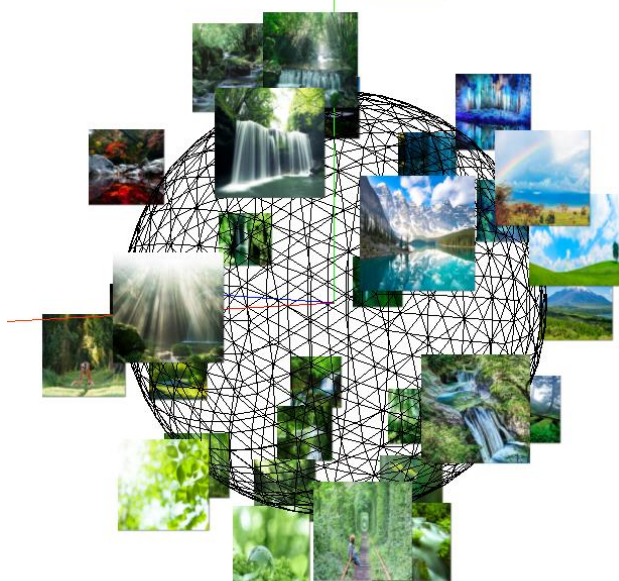


図 8 フィルター学習型球面畳み込み SOM で
自然画像を分類した結果

「世界遺産」と検索した画像を分類した結果、図 9,10 のようになった。「自然」画像のときは配置に大きな変化はなかったが、「世界遺産」画像の配置は、フィルターの影響で大きく変わった。この結果からも、より色の特徴でのまとまりが強くなったクラスタリングがなされていることがわかる。また、同じ建築物でも配置が遠いものがあるため、テクスチャーが捉えられていないことがわかる。

実験結果から、入力画像の平均的な特徴を持ったフィルターになってしまい、テクスチャーが捉えられていないこ

とがわかった。また、入力画像のバリエーションが多いため、特徴がまとまらなかった可能性も考えられる。



図 9 球面畳み込み SOM で世界遺産画像を分類した結果



図 10 フィルター学習型球面畳み込み SOM で
世界遺産画像を分類した結果

6. MNIST 画像を用いた画像分類実験

6.1 実験方法

フィルター学習型球面量み込み SOM の評価を行うために、 28×28 ピクセルの手書きの数字の画像データである MNIST を用いる。フィルター画像の作成には 0~9 の数字の画像 10 枚ずつの計 100 枚を用いた。フィルター作成に用いた画像を図 11 に示す。この画像を 25 分割し、 12×12 ピクセルの画像として学習させ 20 枚のフィルター画像を作成した。分類する画像は、学習とは異なる 0~9 の数字の画像 3 枚ずつの計 30 である。分類する入力画像を図 12 に示す。第 5 章と同様に、フィルター作成用 SOM の競合層は 20 のノードを持ち、クラスタリング結果を示す SOM の競合層は 1280 のノードを持つ。学習回数は 300 回、学習率係数は 0.8、球の半径は 250、近傍領域は 220 と共通している。



図 11 フィルター学習用画像



図 12 分類した入力画像

6.2 実験結果と検討

SOM で学習して得られたフィルターは図 13 の通りである。第 5 章とは違ってグレースケールの画像を用いたため、テクスチャーが捉えられ、曲線も直線も作成できることがわかった。

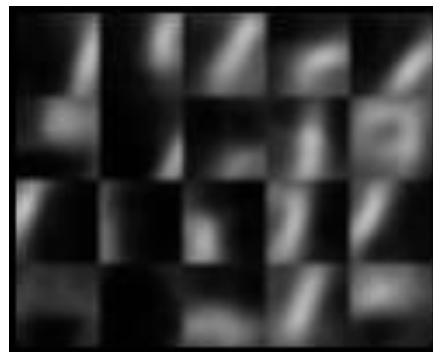


図 13 作成したフィルター画像

30 枚の画像を分類した結果、図 14 のような結果が得られた。曲線や直線の特徴でクラスタリングされていることがわかる。しかし、数字自体は分類できていない。これは、学習に用いた画像が少なく、フィルター画像が大きかったため、細かい特徴の分類ができなかったと考えられる。

今回の評価は感覚的なもので、精度の評価ができていない。クラスタリングの精度確認のために、ラベルデータを持つ画像を分類する必要がある。また、画像の内容を示すラベルデータを学習に用いることで、より精度を高めることができる。

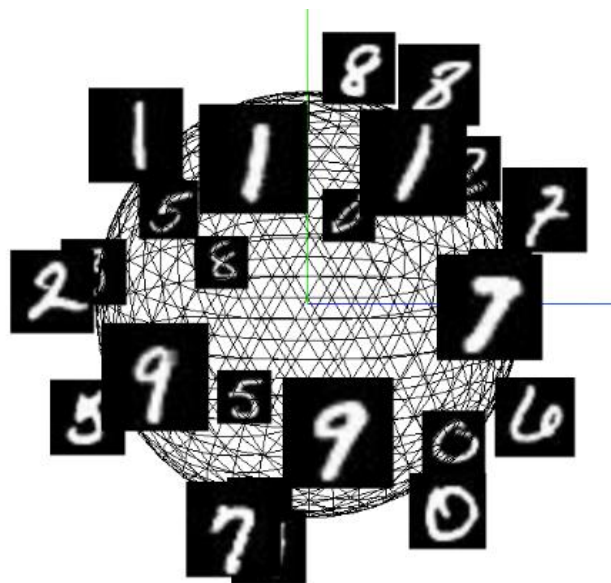


図 14 フィルター学習型球面量み込み SOM で MNIST 画像を分類した結果

7. フィルター学習用 SOM の比較実験

7.1 実験方法

第6章と同じく MNIST を用い、フィルター学習用 SOM の競合層を 20 面と 80 面の 2 パターン準備し比較した。フィルター画像の作成には 0~9 の数字の画像 600 枚ずつを用いた。この画像を 169 分割し、4×4 ピクセルの画像として学習させ 20 枚のフィルター画像を作成した。第6章の研究ではフィルターサイズが大きすぎてはっきりしたテクスチャーが得られなかったため、フィルターサイズを小さくすることにした。

その後、学習とは異なる 0~9 の数字の画像 400 枚ずつの画像データに対してフィルター学習を行う畳み込み球面自己組織化マップで分類した。出力層の参照ベクトルから判断されるラベルと実際に割り当てられたデータの正答率をそれぞれ求める。

7.2 実験結果と検討

SOM で学習して得られたフィルターは図 15,16 の通りである。出力層が多いフィルターからはランダムにフィルター画像を選んでいるため、似たような画像が複数あるような形になってしまった。また、出力層が多いフィルターの方が、ぼんやりとした画像になった。



図 15 SOM の出力層が 20 面のフィルター



図 16 SOM の出力層が 80 面のフィルター

4000 枚の画像を分類した結果、図 17,18 のような結果が得られた。正答率はそれぞれ表 1 のとおりである。フィルター用 SOM の出力層が大きい方が、フィルターのパターンがぼやけたようなものになったため、精度の低下につながったと考えられる。

表 1 フィルター学習用 SOM の競合層の
ノードの数を変えたときの
クラスタリング正答率

	出力層 20	出力層 80
正答率[%]	42.1679	37.5976



図 17 SOM の出力層が 80 面のフィルターでのクラスタリング

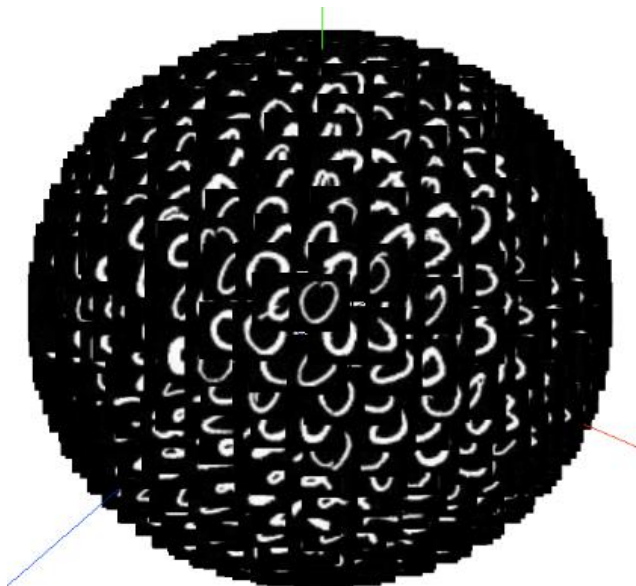


図 18 SOM の出力層が 20 面のフィルターでのクラスタリング

8. おわりに

本研究では球面 SOM で画像を学習してフィルターを作り、球面畳み込み SOM でクラスタリングを行った。第5章の研究結果から、あらかじめ用意しておいたフィルターとは異なるフィルター画像が SOM で学習して得られることがわかった。また、あらかじめ用意しておいたフィルターでの分類と SOM で学習したフィルターでの分類を比較すると、SOM で学習したフィルターの方が、より色の特徴に従った分類がなされることがわかった。しかし、得られたフィルターは連続的に濃度に変化したような画像で、テクスチャーの特徴をとらえてはいなかった。そのため、画像の対象の形による分類はできなかった。また、画像にラベル

データがなかったため、分類が正しいかどうかの評価ができなかった。そこで、第6章では、MNISTという手書きの数字の画像データを分類し、フィルター学習型球面畳み込みSOMを評価した。この実験ではフィルター画像を小さくすることで、テクスチャーへの対応を試みた。また、フィルター学習用のSOMの出力層を20にしたものと80にしたものの2通り試した。出力層が大きいほうがフィルター画像同士の違いが激しくなく、正答率が低かった。一方、出力層が大きいSOMは、ぼやけたような画像を作成してしまうため、取捨選択を行ってフィルターに用いるべきであることがわかった。また、用いたデータが大きすぎたために学習がうまくいかなかった可能性があるため、パラメーターの見直しが必要である。

正答率は高くなかったものの、畳み込み用のフィルターを工夫することでクラスタリング精度の向上につながることもわかった。そのため、今後の課題としては、画像をグレースケールで処理してテクスチャーをとらえたフィルターを作成するといった、他のフィルターの作成方法を試す必要がある。また、SOMで分類する前の新たな画像処理の方法として、オートエンコーダというデータを圧縮した情報を球面SOMでクラスタリングを行うという方法が考えられる。

謝辞

本研究について御助言や御協力いただいた本研究室の教授、学生、大学院生の皆様に謹んで感謝の意を表する。

参考文献

- [1] 萩原将文：ニューロ・ファジィ・遺伝的アルゴリズム，1994
- [2] 畳み込みネットワークの「基礎の基礎」を理解する～ディープラーニング入門 | 第2回，
<https://www.imagazine.co.jp/%E7%95%B3%E3%81%BF%E8%BE%BC%E3%81%BF%E3%83%8D%E3%83%83%E3%83%88%E3%83%AF%E3%83%BC%E3%82%AF%E3%81%AE%E3%80%8C%E5%9F%BA%E7%A4%8E%E3%81%AE%E5%9F%BA%E7%A4%8E%E3%80%8D%E3%82%92%E7%90%86%E8%A7%A3%E3%81%99/>
- [3] 豊住健太，堂菌 浩：球面畳み込み自己組織化マップを用いた画像のクラスタリング，情報処理学会第81回全国大会論文集，2019年3月
- [4] Hiroshi Dozono ; Gen Niina ; Satoru Arak: Convolutional Self Organizing Map, Proceedings of 2016 International Conference on Computational Science and Computational Intelligence (CSCI)., IEEE Xplore Digital Library, Dec,2016