

# 表現分離情報を用いたびまん性肺疾患識別用 Deep Convolutional Neural Network のアーキテクチャ設計

鈴木 聡志<sup>1,a)</sup> 飯田 のどか<sup>1</sup> 庄野 逸<sup>1</sup> 木戸 尚治<sup>2</sup>

**概要:** Deep Convolutional Neural Network (DCNN) は階層型ニューラルネットワークモデルの一種であり、近年、Krizhevsky らが DCNN を用いて大規模自然画像識別コンテストで state-of-the-art を記録するなど [6]、画像識別や音声認識タスクで従来手法を凌駕する精度を示し、注目を集めている。しかし、明確な設計指針が存在しない為 DCNN の精度を引き出すためのアーキテクチャ設計の手法については多く議論がなされておらず、不十分な点が多い。本研究では、線形 Support Vector Machine (SVM) を用いて、各階層での識別するクラスの分散情報を評価しこの情報を基に DCNN のアーキテクチャ設計を行う手法を提案する。我々は従来の DCNN に対してこの手法を適応し、びまん性肺疾患と呼ばれる疾患の X 線 CT 画像に適合するようアーキテクチャの修正設計を試みた [9], [10]。その結果、修正を行う前の DCNN よりも高い識別精度を示す事に成功した。

**キーワード:** 深層畳み込み神経回路, 線形サポートベクターマシン, びまん性肺疾患画像解析

## Architecture Design of Deep Convolutional Neural Network for Diffuse Lung Disease Using Representation Separation Information

SATOSHI SUZUKI<sup>1,a)</sup> NODOKA IIDA<sup>1</sup> HAYARU SHOUNO<sup>1</sup> SHOJI KIDO<sup>2</sup>

**Abstract:** In this work, we propose a new architecture design of Deep Convolutional Neural Network (DCNN) with representation separation information of intermediate layers. The DCNN is one of the multi-layer neural network models. In recent years, the DCNN is attracting attention by its state-of-the-art performance in the image and speech recognition tasks. For example, Krizhevsky *et al.* showed the state-of-the-art performance in the large scale image recognition in 2012 [6]. However, the design for the architecture of the DCNN has not been discussed much since we have not found effective guideline to construct. We try to modify the architecture of the DCNN for the Diffuse Lung Disease (DLD) image classification task [9], [10], and confirm that the modified DCNN shows better performance than that of the original one.

**Keywords:** Deep Convolutional Neural Network, Linear Support Vector Machine, Diffuse Lung Disease image analysis

### 1. まえがき

Krizhevsky らが提唱した AlexNet などの Deep Convolutional Neural Network (DCNN) モデルによって昨今の画像識

別タスクの精度は飛躍的に向上した [3], [6]。DCNN モデルは 1980 年代に Fukushima が提唱した Neocognitron と同様の構造を持ち、入力から自動で特徴抽出器を構築することが可能なモデルである [4], [7]。代表的なモデルである AlexNet は ImageNet データセットのような大規模データセットの識別タスクで他手法を凌駕する性能を示したが、モデルが複雑であるためネットワーク内部の挙動や処理の解析や、学習過程でどのようにして高い識別精度を得るか

<sup>1</sup> 電気通信大学大学院情報理工学研究所, 〒182-8585 東京都調布市調布ヶ丘 1-5-1

<sup>2</sup> 山口大学大学院応用医工学研究科 〒755-8611 宇部市常磐台 2-16-1

<sup>a)</sup> sat.suzuki@uec.ac.jp

という考察が不十分であった [2], [6], [11]. 加えて, DCNN には多くのハイパーパラメータが存在するため, 対象データに依存したアーキテクチャ設計が困難という問題点がある [11]. 先行研究において我々は DCNN の階層ごとの特徴表現を解析するため Support Vector Machine (SVM) ヒストグラムを提案し, DCNN を十分なデータサンプルで学習すると, SVM ヒストグラムの形状が階層を経る毎に細くなる現象を発見した [9]. 本研究では, この事実を基に SVM ヒストグラムのクラス内分散に着目し, びまん性肺疾患 [9], [10] と呼ばれる疾患の X 線 CT 画像に適合する DCNN の設計を試みた.

### 1.1 びまん性肺疾患

びまん性肺疾患とは, 肺の広範囲にわたって異常な陰影が観察される病気の総称である. びまん性肺疾患の陰影は, 淡い陰影パターンを呈し, 広範囲にわたって観察されることが多く, 陰影パターンが複雑多様である事が知られている [10]. また, 医療用 CT 画像はラベル付けに医師の診断が必要という特性上, 大量のデータを取得する事ができない為, 自然画像データセットに用いるような大規模 DCNN には直接適用する事が難しい事が知られている [9].

## 2. 手法

### 2.1 Deep Convolutional Neural Network

本研究で用いる Deep Convolutional Neural Network (DCNN) は Fukushima の提唱した Neocognitron と同様な構造を持ち, 階層間の演算が畳み込み演算で記述できるのが特徴である. 本研究では, 教師付き学習の一種である誤差逆伝搬 (BP) 法で動作する一般的な DCNN を用いる [7]. DCNN モデルは一般に 2 次元カラー画像の入力から階層を複数回経て各クラスの分類確率をベクトルで出力する様なシステムとなっている.

それぞれの階層は, (i) 前段の出力に対して畳み込みを行う convolution 層 (ii) 活性化関数 Rectified Linear (ReLU) 関数 ( $relu(x) = \max(x, 0)$ ), (iv) [optionally] 局所正規化処理, そして (iii) [optionally] 空間プーリング処理, から成る. 図 1 に先行研究で用いた CaffeNet のアーキテクチャを示す. この DCNN は, AlexNet と類似のアーキテクチャを持つ [6].

### 2.2 線形 Support Vector Machine を用いた分離度評価

本節では, DCNN の中間層における特徴表現を可視化する手法として著者らが先行研究で提案した SVM ヒストグラムについて説明する [9]. 線形 SVM は特徴空間における入力データ  $x$  に対して,  $1/\|w\| \min_n [t_n(w^T \phi(x_n) + b)]$  を最大化するような決定境界  $y(x) = w^T \phi(x) + b = 0$  を作成する. なお,  $\phi(x)$  は入力パターン  $x$  に対する中間層の表現,  $t$  は教師信号,  $n$  は入力データのインデックスをそれぞれ示す.

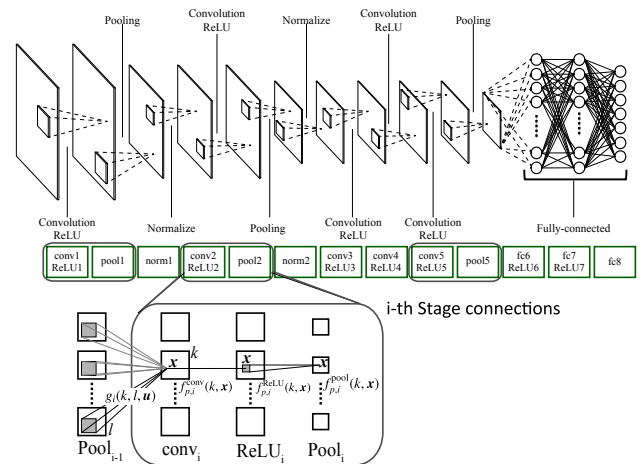


図 1 DCNN モデルの一種である, CaffeNet の概要 (図中上部) と処理の詳細 (図中下部). DCNN モデルは一般的に図中下部のような Convolution 層 → ReLU 層 → Pooling 層の処理を繰り返す事で特徴表現を獲得する.

先行研究において我々は DCNN の各階層から特徴表現を抽出し, 得られた表現を用いて学習した決定境界  $y(x) = w^T \phi(x) + b = 0$  とテストデータの特徴表現  $\phi(x)$  との距離を分離度の指標とし, クラスに対する射影  $\{y(x_n)\}$  をヒストグラムとして解析する手法を提案した [9]. 以下では, このヒストグラムを“SVM ヒストグラム”と呼ぶ.

## 3. 実験

### 3.1 実験データ

本研究で使用したデータは大阪大学の提供による, びまん性肺疾患を対象とした X 線 CT 画像であり, 医師の指導の下, 7つのクラスに分類されている. まず, びまん性肺疾患を呈する陰影は 4 クラスで, それぞれ浸潤影 (CON), すりガラス状陰影 (GGO), 蜂巣状陰影 (HCM), 網状影 (RET) である. びまん性肺疾患以外の陰影は 2 クラスで, びまん性肺疾患と似た陰影を呈する肺気腫 (EMP) と, 腫瘍などが呈する粒状影 (NOD) である. さらに, 正常状態 (NOR) を加えた合計 7 クラスで分類を行う. CT 画像から医師の指定した病変部分を  $32 \times 32$  [pixel] の大きさのパッチ画像として切り出した. これを ROI (Region of Interest) 画像と呼ぶ. 病変部は必ずしも  $32 \times 32$  [pixel] の正方形領域で切り出せるわけではないため, 領域内に病変部が 80% 以上含まれるものを用いた. 各陰影パターンの典型例を図 2 に, CT 画像から切り出した ROI 画像の DCNN 特徴学習用のデータと, 識別実験における識別器の訓練用と評価用に分けた内訳を表 1 に示す.

また, 表 1 において, “DCNN 特徴学習用” と “SVM 学習と評価用” の ROI 画像の枚数のオーダーが異なっているが, これは ROI 画像の採取の仕方に依存している. “DCNN 特徴学習用” データでは, DCNN の学習に用いる事を前提に ROI 画像の抽出の際, 画像の重なりを許容し, 8 [pixel] ず

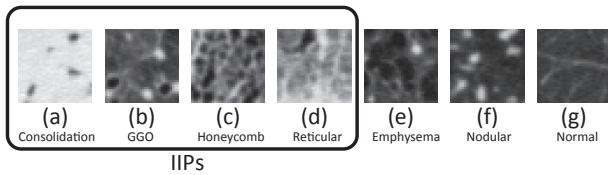


図2 ROI画像7種類の典型例。a, b, c, dがびまん性肺疾患の陰影クラスである。

表1 びまん性肺疾患のX線CT画像から抽出したROI画像の学習・評価用の枚数内訳。

	Dataset 1			Dataset 2		
	Patients	Train	Val.	Patients	Train	Val.
CON	13	143	26	14	247	16
GGO	14	609	46	14	443	53
HCM	10	282	73	9	782	32
RET	8	210	66	9	210	66
EMP	10	4406	296	11	3760	296
NOD	9	762	65	10	591	85
NOR	11	5371	355	11	4595	406
Total	—	11783	927	—	10628	916

つラスタスキャンしながらROI画像を得ているのに対し、“Validation”データでは重複を許容しないように抽出を行っているため抽出元のCT画像の枚数は同数程度であるが、ROI画像の枚数のオーダーが異なっている。また、それぞれのROI画像の抽出元のCT画像の振り分けは医師の指導に従った。

## 4. 実験

表1のDataset1のTrainデータを用いて学習したCaffeNetにおいて、Dataset2のVal.データの特徴表現を抽出し、2.2節のSVMヒストグラムを作成する。

### 4.1 結果と評価

7クラス分類の全ての組み合わせ  ${}^7C_2 = 21$  通りのSVMヒストグラムを作成し、それぞれの階層のクラス内分散を図3にまとめた。Conv1層からPool2層まではクラス内分散が低い水準で推移しているが、Conv3層以降ではクラス内分散が大きくなっている事が分かる。これは、CaffeNetが自然画像に対してチューニングしたため、びまん性肺疾患画像に対しては必要以上に階層が深いアーキテクチャとなっていることを示唆している。

## 5. 提案手法と精度評価

以上の議論から、CaffeNetをびまん性肺疾患に適合したアーキテクチャにするため、従来のCaffeNetから図3からクラス内分散が大きくなってしまっているConv3,4層を取り除いた新たなアーキテクチャを提案し、従来のCaffeNetとの比較を行っていく。

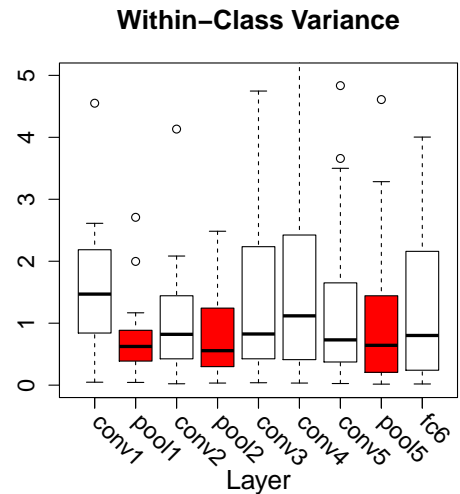


図3 SVMヒストグラムの全ての組み合わせにおけるクラス内分散の箱ひげ図。Conv3層以上でクラス内分散が大きくなっている事が分かる。

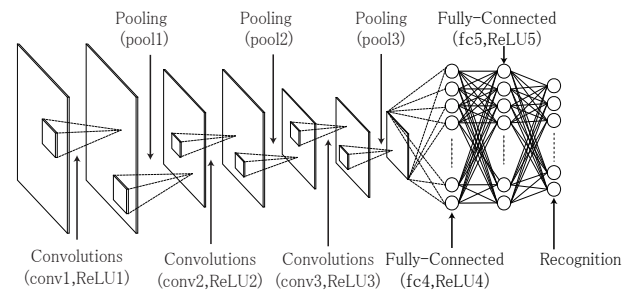


図4 CaffeNetからconv3,4層を取り除いた新たなDCNNアーキテクチャの概要。なお、畳み込みフィルタの枚数や全結合層のニューロンの数などは変更していない。

表2 CaffeNetと我々の提案したアーキテクチャとの精度比較。評価指標には2-fold CV法を用いている。

DCNN	Accuracy
CaffeNet (5-conv + 3-full connect)	85.26 ± 2.16 [%]
Ours (3-conv + 3-full connect)	88.39 ± 0.01 [%]

### 5.1 精度評価

3.1節で用いたびまん性肺疾患画像データを用いて学習と評価を行う事で、通常のCaffeNetと我々が提案したDCNNの2つのアーキテクチャの精度評価を行い、結果を表2にまとめた。また、評価指標は表1で分割した2つのデータセットによる2-fold Cross Validation (CV)法を用いた。表2から、我々の提案したDCNNアーキテクチャがCaffeNetよりも高い識別精度を示しており、階層数の改善に成功している事が分かる。

### 5.2 SVMヒストグラムを用いた分離度解析

設計指針として取り入れたSVMヒストグラムのクラス内分散が設計指針通り小さくなっているのか、調査を行う。図5は我々が新たに提案したモデルにおけるSVMヒストグラムのクラス内分散である。従来のCaffeNetを用いて

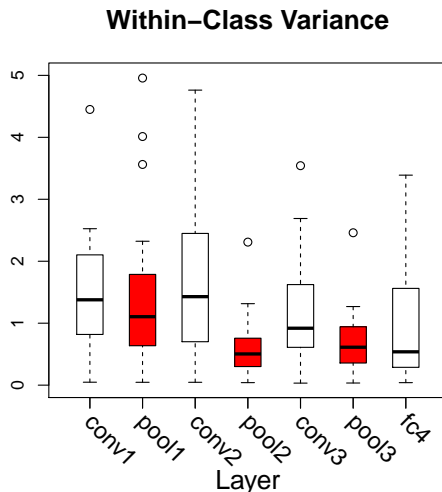


図5 新たなアーキテクチャにおける SVM ヒストグラムのクラス内分散. 従来の CaffeNet よりもクラス内分散が小さくなっていることが分かる.

いる, 図3 よりもクラス内分散が低くなっており, 新たなアーキテクチャが設計指針通りに学習を行っている事が示唆される.

## 6. 結論

本研究では, DCNN のアーキテクチャのチューニング手法として SVM ヒストグラムを用いたクラス分離度を用いる手法を提案した. その結果, 従来の CaffeNet をびまん性肺疾患画像の識別用にチューニングし精度向上をもたらす事に成功した.

びまん性肺疾患画像のようなデータの多量に用意できないデータベースにおいて, DCNN を用いる手法は先行研究 [9] のように自然画像を用いた転移学習手法などが提案されていた. しかし, 先行研究で用いた手法は転移学習に2次元の自然画像を用いているという点から, びまん性肺疾患画像のような3次元画像への拡張が困難であった. 本研究の手法では転移学習手法を用いないため, 3次元画像への拡張も可能だと考えられる.

今後の課題としては, 先行研究 [9] で示したような水準まで, SVM ヒストグラムのクラス内分散を減少させることである. これは, フィルタのサイズなど階層数以外のハイパーパラメータをびまん性肺疾患画像用にチューニングする事で可能になると考えられる.

謝辞 本研究の一部は, JSPS 科研費 26120515 及び 16H01542 の助成を受けたものである. また, 貴重なデータを提供していただいた, 大阪大学附属病院に深く感謝いたします.

## 参考文献

[1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[3] L. Deng and D. Yu. Deep learning: Methods and applications. Technical Report MSR-TR-2014-21, Microsoft Research, May 2014.

[4] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, Vol. 36, No. 4, pp. 193–202, 1980.

[5] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. B. Girshick, S. Guadarrama, and T. Darrell. Caffe: Convolutional architecture for fast feature embedding. *CoRR*, Vol. abs/1408.5093, , 2014.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pp. 1097–1105. Curran Associates, Inc., 2012.

[7] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L.D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, Vol. 1, No. 4, pp. 541–551, 1989.

[8] D. Scherer, A. Müller, and S. Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. In *ICANN'10 Proceedings of the 20th international conference on Artificial neural networks: Part III*, Vol. 6354, pp. 92–101, 2012.

[9] H. Shouno, S. Suzuki, and S. Kido. A transfer learning method with deep convolutional neural network for diffuse lung disease classification. In *Neural Information Processing, 22nd International Conference, ICONIP 2015, Istanbul, Turkey, November 9-12, 2015, Proceedings, Part I*, Vol. 9489 of *Lecture Notes in Computer Science*, pp. 199–207. Springer, 2015.

[10] R. Xu, Y. Hirano, R. Tachibana, and S. Kido. Classification of diffuse lung disease patterns on high-resolution computed tomography by a bag of words approach. In *MICCAI*, Vol. 14, pp. 183–190. Springer-Verlag Berlin Heidelberg, 2011.

[11] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. In *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I*, pp. 818–833, 2014.