I-2 講演会

ディープラーニングを用いた画像テクスチャ解析 -データ駆動科学への橋渡しを目指して-

講師:電気通信大学 大学院情報理工学研究科 電気通信大学 人工知能先端研究センター 教授 庄野 逸

開催日 : 2019年2月18日

開催場所 : リロの会議室(コンフォート田町) J会議室

参加者 : 23 名

記 : 豊吉 直樹 *

1. はじめに

ここ数年、「第三次 AI ブームの到来」と呼ばれるほど、AI (Artificial Intelligence) への関心が高まってきており、協会の会員各社においても、情報機器分野での画像処理・画像認識・自然言語処理、ヘルスケア分野での医療画像処理などに AI の導入が進んでいる。

国立大学法人電気通信大学は、国立大学では初の研究センターとなる「人工知能先端研究センター (Artificial Intelligence eXploration Research Center: AIX)」を2016年7月1日に設立した。AIXは「AI for X」を基本理念とし、「科学のためのAI」、「デザインのためのAI」、「サービスのためのAI」の三つの軸を掲げ、それらの共通基盤となる「汎用人工知能(汎用AI)」の実現を主眼とする新たな人工知能研究を推進することを目的としている。

取り組む技術分野は、ディープラーニング、機械学習、スパースモデリング、高次元データ駆動科学、全脳アーキテクチャ、脳神経科学、自然言語・画像・音

声処理、ロボティクスなど多岐に渡るが、今回は会員 各社の事業にも関連が深い画像関連のテーマを選定さ せていただき、ディープラーニングを用いた画像テク スチャ解析と題してご講演いただいた。

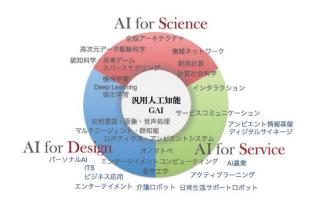


Fig. 1 AIX の取り組む技術分野 (HP より)

2. 講演内容

2.1. 画像認識解法の変遷と課題

最初にディープラーニングの簡単な説明があった 後、第1世代ニューラルネット (NN) 、第2世代

^{*} 技術調査専門委員会委員

NN、ディープラーニング以前の機械学習から深層畳み 込みニューラルネット (DCNN) に至る画像認識解法の 変遷と課題が説明された。



Fig. 2 講演風景(冒頭)

ディープラーニング以前の画像認識では、画像特性 (エッジ特徴等)の特徴量構築と機械学習を浅いネットワークで行っていたが、ハンドデザインによる特徴 抽出機構設計の難しさが顕在化し、機械学習による表 現獲得に移行している。

多層パーセプトロンと誤差逆伝搬法を用いる第2世代NNには、パラメータ数が過多になりがちで、深い階層だと勾配情報が拡散する、全体のトレーニングが難しいなどの問題点があり、問題解決へのアプローチのひとつとして、DCNNの概観が解説された。



Fig. 3 講演風景 (DCNN の概観)

DCNN は生物の視覚系を模倣した階層型ニューラル

ネットで、画像処理分野では既にデファクトスタンダードとなっており、特徴抽出部と識別部の結合重みをデータに適合させられる特徴を持つ。

2.2. 医用画像への DCNN 適用

一般に NN を訓練するには十分な数の学習データが必要となるが、医用画像では大規模データセットが獲得しにくいという課題がある。庄野教授の研究室では、少ない学習データでかつテクスチャ構造を持つ陰影パターンに適用ができる学習法が研究されており、DCNNに転移学習を用いられている。転移学習とは、目的とは違う問題で得た知識を転用する手法である。

NN は特徴抽出層と識別層から構成されているが、 DCNN における転移学習では、識別層を付け替え、特徴 抽出層を転用して異なるデータセットでネットワーク 全体を再学習させることで、特徴抽出層の性能向上を 図っている。

自然画像のデータセットを用いた転移学習により、 ターゲットとしているびまん性肺疾患陰影の識別性能 が向上していることが示された。さらに自然画像によ り画像一般の特徴を獲得させ、テクスチャ画像により 陰影パターン(テクスチャ)に有用な特徴を獲得させ ることを目的に、自然画像とテクスチャ画像を用いた 2 段階に渡る転移学習への拡張が行われている。

2 段階転移学習を行なったモデルは、それぞれの画像データセット単独で転移学習を行ったモデルと比較して最も高い識別性能を示していた。またこの2段階学習を行ったモデルは、少ないデータセットに対する頑健性でも優れていることが示された。

2.3. メカニズムの解明に向けて

このように2段階転移学習が識別性能の向上に効果 的であることは分かったが、DCNNが画像に対する特徴 表現をどのように学習するのかといった転移学習のメ カニズムの解明も併せて行われている。

通常、ディープラーニングの中味はブラックボック スとして扱われている事が多いが、内部で何が起きて いるのかといった解析にまで踏み込まれて研究が行わ れており、非常に興味深いお話だった。

信号が通るパスを再現するバックプロジェクション (選択的注意機構モデル)を導入し、中間層の表現を 直接見る手法を用いて、DCNNが画像の何処を、何を観 ているのか、DCNNの階層を経るごとにどのような表現 になるのかを研究されている。異なる画像セットで学 習した DCNN の特徴表現の違いを観察することで、な ぜ転移学習によって識別性能が向上するのかが可視化 されていた。

自然画像で学習したモデルは入力画像のエッジ構造に活性を示し、テクスチャ画像で学習したモデルは模様に活性を示すようになっており、2 段階で段階的に両方を学習したモデルはそれらの両方に活性を示すようになっていることが説明された。DCNN は学習に用いたデータセットに応じた特徴表現を獲得しており、これはまた同時にデータセットの選択の重要性も示唆しているとのことであった。

データセットは自然画像とテクスチャ画像という全 く異なる種類のものだが、これらを段階的に学習させ ることによって、両方の特徴が識別できようになって きており、これは人間が学習する過程、成長する過程 を見ているようで驚かされた。

3. おわりに

ディープラーニングの基礎から、最新の研究成果までを分かりやすく講演していただいた。

講演会には会員各社から多く方が参加されて活発な 質疑応答があり、聴講者の関心の高さが伺えた。我々 の業界でも多角化が進んでおり、医療イメージングに 取り組んでいる企業も多く、大変参考になるご講演内 容だったと考えている。

また今回は、医用画像への適用を事例とした講演内容だったが、用いられている2段階転移学習の手法は、さまざまな画像認識や画像解析に適用できる可能性があると感じられた。

最後に、庄野教授にはお忙しい中、時間を割いていただき、また非常にわかりやすい講演を行っていただいた。この場を借りて御礼申し上げる。

[添付資料に関するお願い]

理解を助けるため、庄野教授のご厚意により、講演 会の説明資料を添付しています。著作権は庄野教授が 有しておられますので、2次利用を希望される方は直 接、庄野教授まで使用許諾のご相談をお願いします。

ディープラーニングを用いた 画像テクスチャ解析

- データ駆動科学への橋渡しを目指して -

2019/02/18

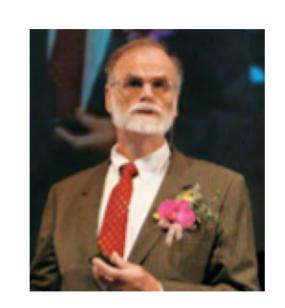
電気通信大学 大学院情報理工学研究科電気通信大学 人工知能先端研究センター

庄野逸: shouno@uec.ac.jp



データ (駆動) 科学とは?

- ☑Jim Gray (1944~2012) が提唱
 - ▼ 第1のパラダイム: 経験科学 (数千年前~ アリストテレス)



- ▼ 第2のパラダイム: 理論科学 (紙と鉛筆, 計算尺) (数百年前~ ライプニッツ)
- ▼ 第3のパラダイム: 情報科学 (計算機) (数十年前~ フォン・ノイマン)
- ☑ 第4のパラダイム: データ科学 (AI?)



Science 2011



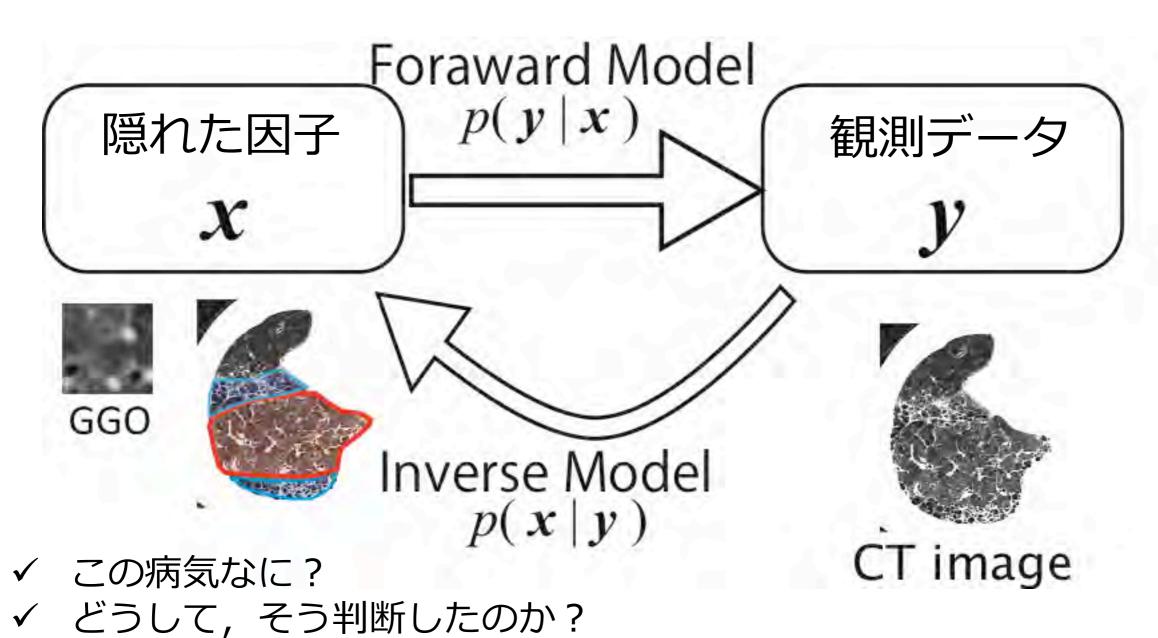
Nature 2008

modified from 岡田真人(東大)

庄野研の向かっている方向(1)

▼Forward/Inverse モデルアプローチ

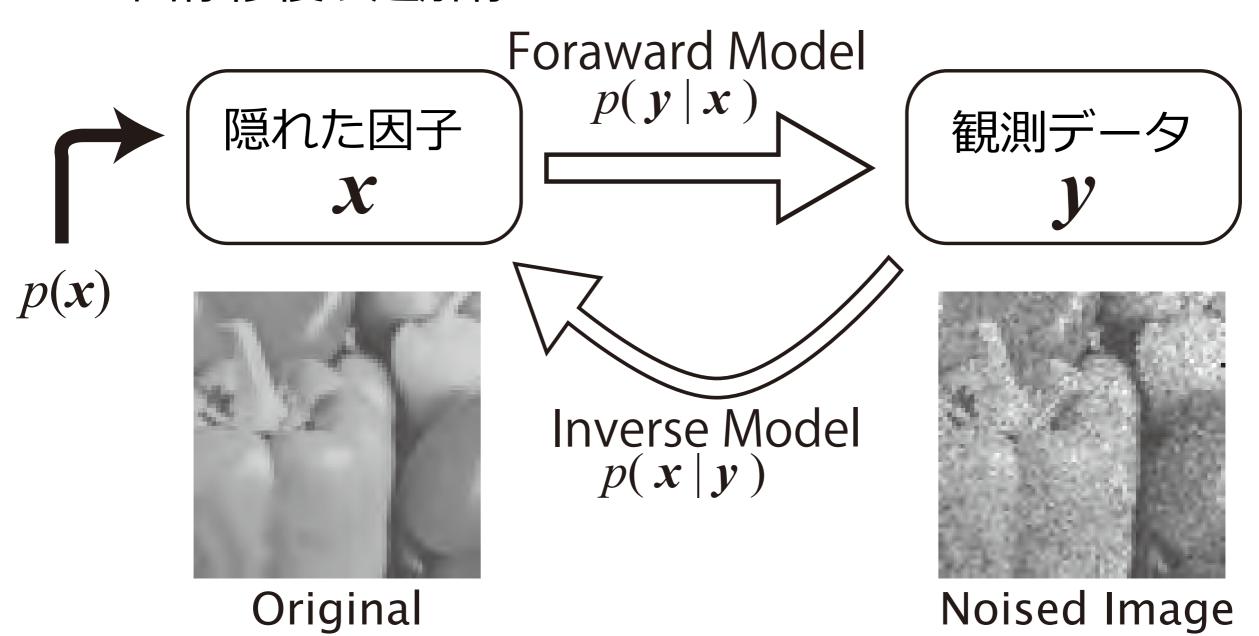
☑判別などのタスク



庄野研の向かっている方向(2)

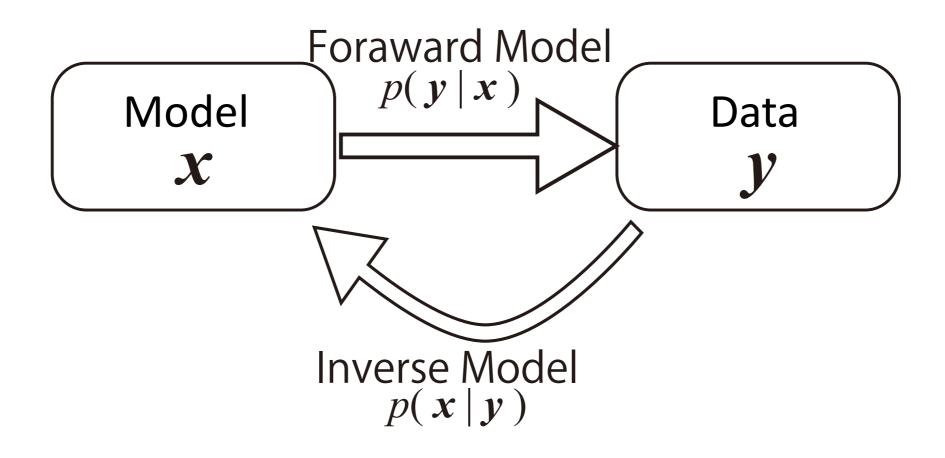
Forward/Inverse モデルアプローチ

画像修復や超解像



2019/02/12 JBMIA 技術調査講演会

庄野研で使うキーテクノロジー

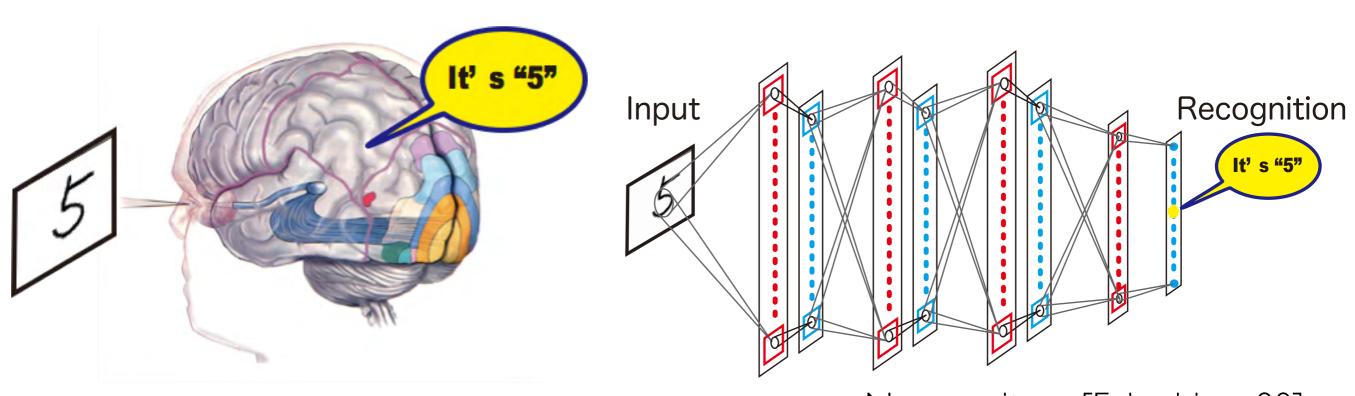


- ▼ Forward モデル
 - ☑ 階層ベイズモデル
 - ▼ マルコフチェーンモンテ カルロ法

- ☑ Inverse モデル
 - ☑ スパースモデリング
 - ☑ マルコフ確率場
 - ☑ 機械学習,深層学習

ディープラーニング(深層学習)とは?

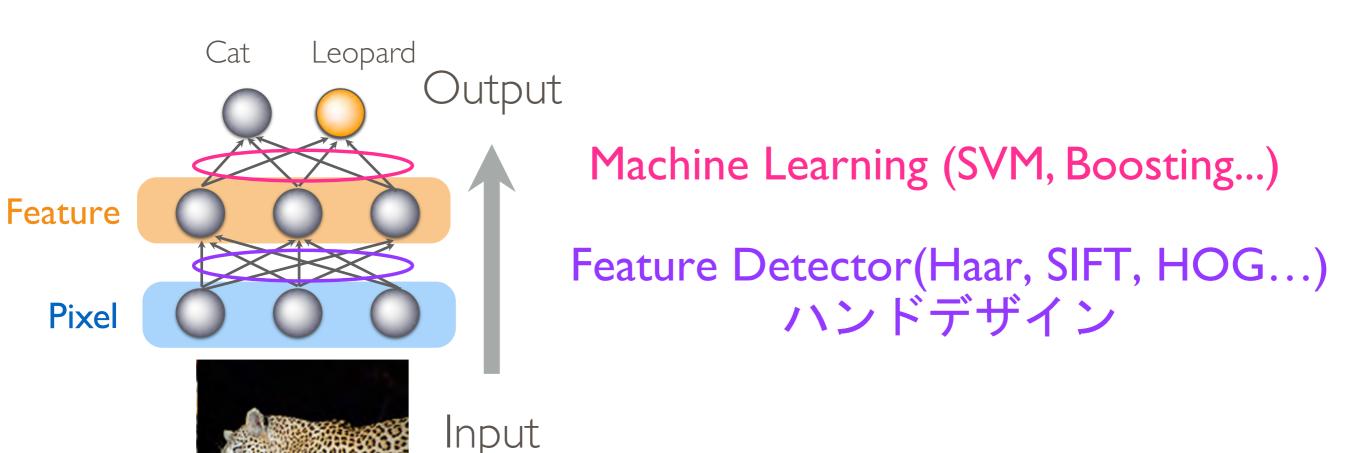
- ☑神経回路(ニューラルネット)モデルを用いた人工知能技術
 - ☑ 脳の働きを模した構造と学習方式
- ☑深い階層構造を持つことが特徴



Neocognitron [Fukushima 80]

ディープラーニング以前の画像認識課題解法

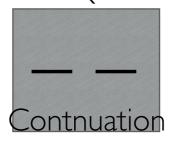
☑画像特性(エッジ特徴等)の特徴量構築+機械学習 ☑浅いネットワーク(Shallow Network model)

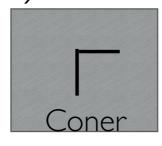


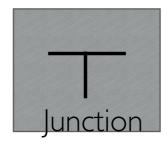
特徴抽出器設計の難しさが顕在化 誤差低減率が 1 [%/years]

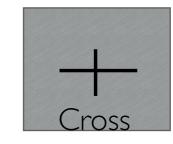
特徴抽出機構の設計

☑どうやって(中程度に複雑な)特徴検出器を作るか?

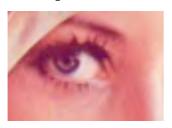








Object parts:







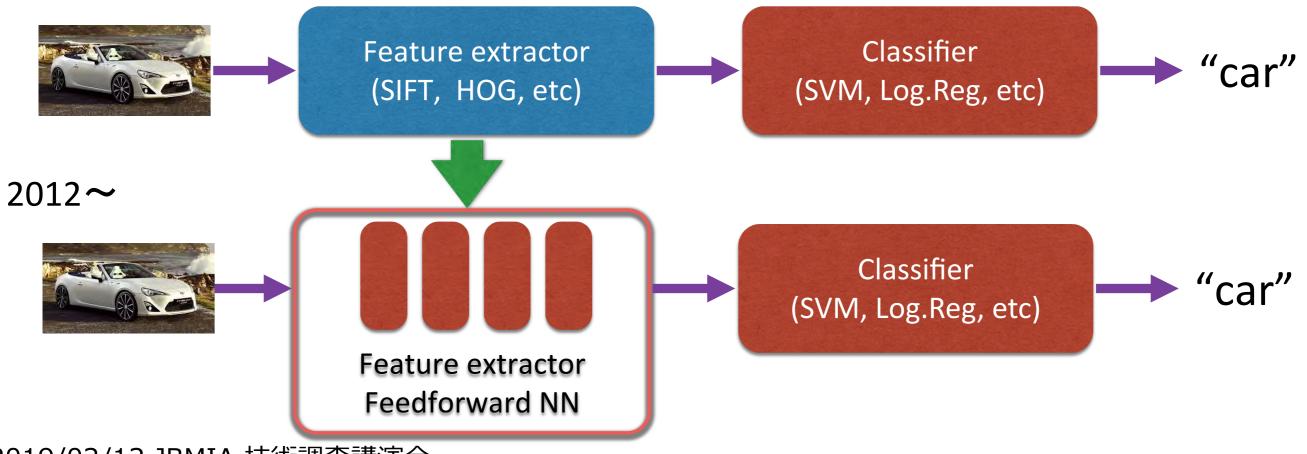


✓ハンドメイドな特徴量はしんどい →機械学習による表現獲得

Computer Vision 業界でのディープラーニング

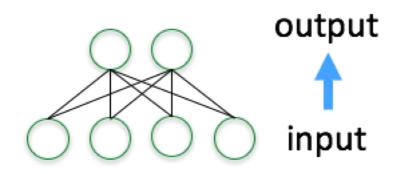
☑Computer Vision 業界での特徴抽出の変遷

☑ディープラーニング以降(2012~) 特徴表現の学習 + 学習可能な識別器



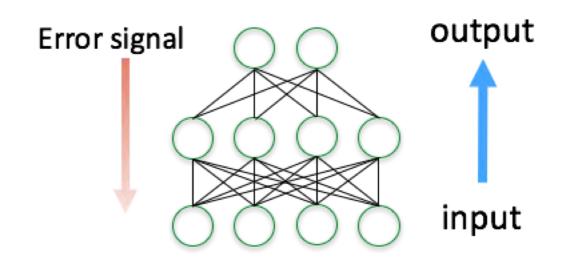
NNの簡単なまとめ

- ☞ 第1世代 NN (1960~1970)
 - ☆ ネットワーク構造: (単純) Perceptron
 - ▼ 学習: 誤り訂正学習



線形分離問題しか 解決できない

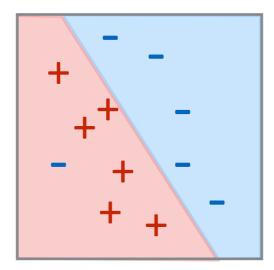
- 第2世代 NN (1980~1990)
 - ネットワーク構造:多層 Perceptron (MLP)
 - ☞ 学習: 誤差逆伝播法 (BP)



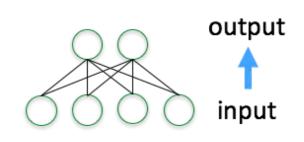
#パラメータ数 » #学習データ

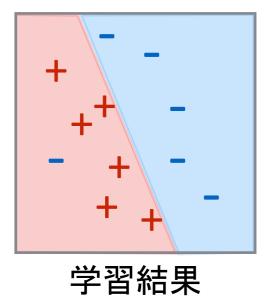
- ☑ 過学習し易い

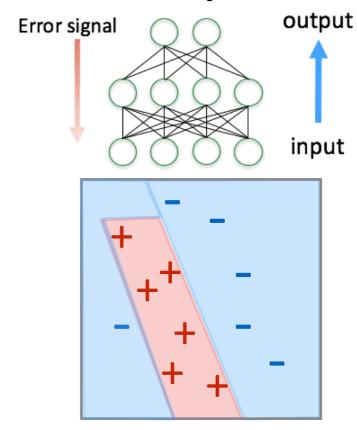
NNの簡単なまとめ: cont'd



training samples



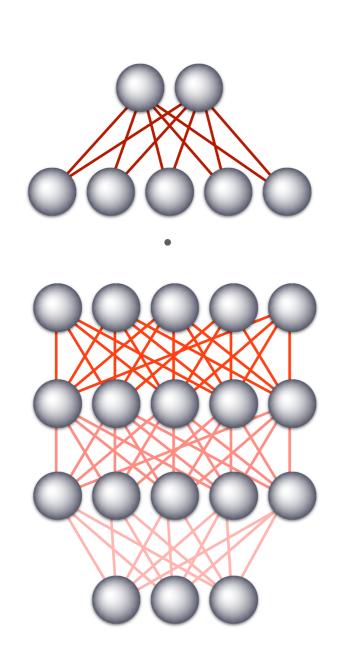




学習結果

#パラメータ数 >> #学習データ 分離境界が必要以上に複雑に...

第2世代 NN の問題点



- ☑データに対してパラメータ数が過多!
 になりがち
 - ☑ 過学習問題 (訓練誤差 » 汎化誤差)
- ☑ 深い階層だと勾配情報が拡散する
 - ☑ 全体のトレーニングは難しい
 - ☑ 全結合型 NN で顕著

問題解決へのアプローチ

「階層が深いモデルを諦めたほうがいいんでは?」派 浅いネットワークで頑張る→ SVM, Boosting, etc…

「勾配さえ生き残ってくれればなんとかなる」派

勾配消失の補償→ ReLU ユニット / LSTM モジュール

「学習を各階層でちょっとづつ」派

階層ごとの教師なし学習の導入→RBM, クラスタリング, Sparse Coding, etc

「最適化を頑張る」派

鞍点からの脱出に学習レートを弄る→ AdaGrad, AdaDelta, Adam, etc…

「(適切に) 自由度削れば良い」派

アーキテクチャに対する制約→ネオコグニトロン, 深層畳み込みニューラルネット (DCNN)

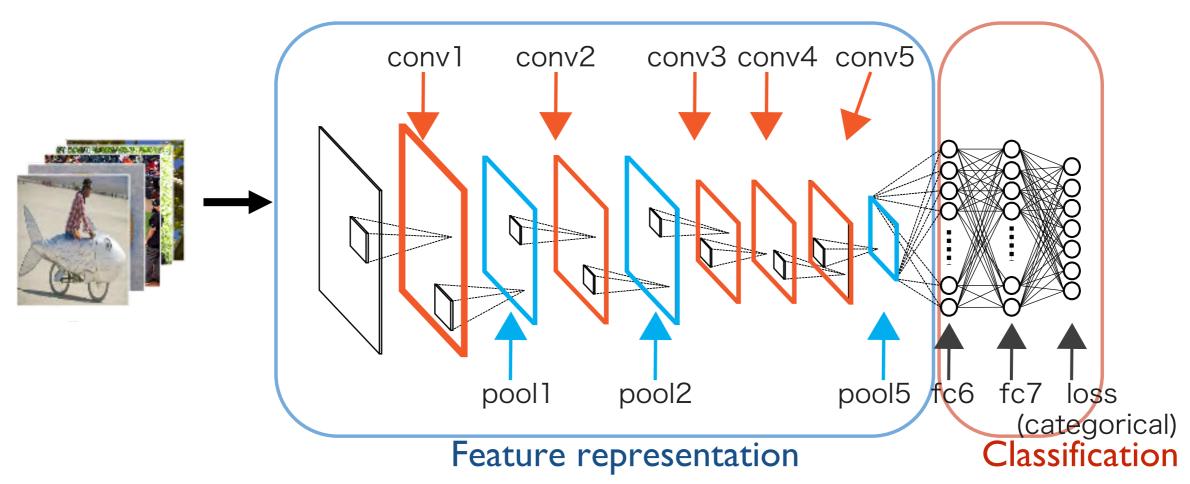
学習に対する制約→Sparse Coding, Drop out

DCNNモデル概観

生物の視覚系を模倣した階層型ニューラルネット

画像処理分野では,既にデファクトスタンダード(AlexNet, VGGなど)[Krizhevsky+12][Simonyan+15]

特徴抽出部と識別部の結合重みをデータに適合させられる

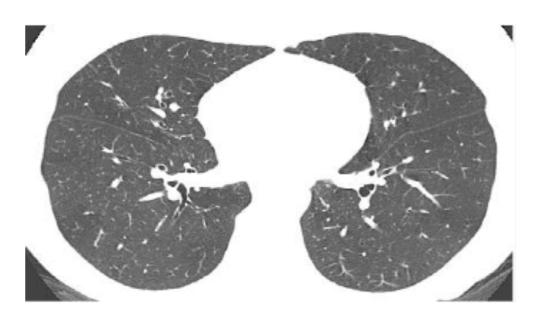


医用画像へのDCNN適用

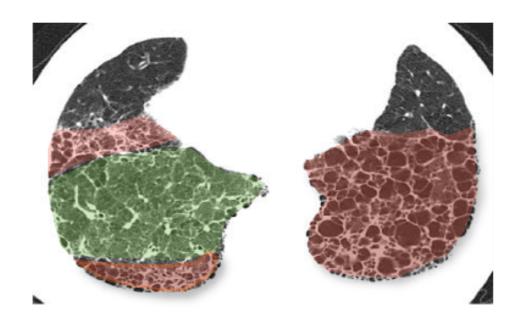
情報処理学会論文誌数理モデル化と応用(TOM), 11(3),74-83 (2018-12-20), 1882-7780

びまん性肺疾患とその肺CT画像概観

- ☑肺広範囲に病巣が広がる疾患の総称
 - ☑肺の線維化を引き起こすことによる呼吸不全
- ☑早期発見,加療が重要
- ☑Diffuse Lung Disease (DLD) パターンから分類したい ☑形状ではなく陰影パターン(テクスチャ)の識別



Normal Lungs



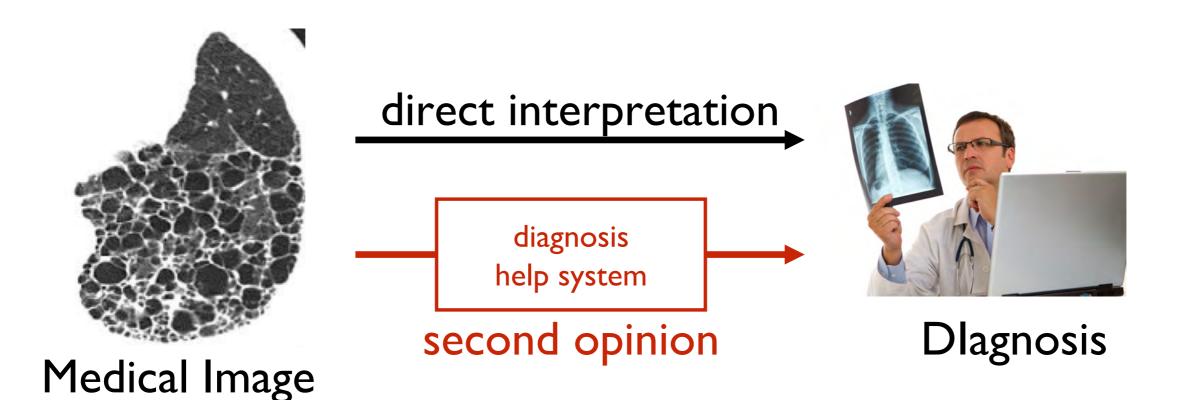
IIP Lungs (HCM, GGO)

計算機支援診断システム構築の目標

専門医の不足,診断スキルのばらつき

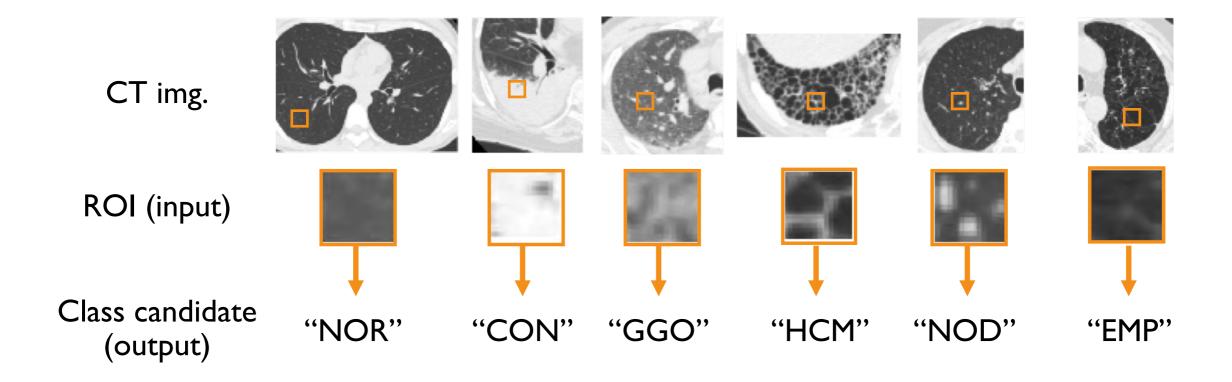
CT画像の高精細化, セカンドオピニオン制度による負担の増大

計算機支援診断(CAD)システム導入による負担軽減



パターン認識技術によるCADシステム

- ☑DLD ROI (region of interest) パターン判別機構の確立 ☑従来手法: 局所特徴量と識別器の組みあわせ [Gangeh+10][Sorensen+10][Xu+11]
- ☑大規模データセットは獲得しにくい 症例数で10² オーダーが限度



びまん性肺疾患データセット

	Training	Evaluation	
Conolication(CON)	143(13patients)	26(14patients)	
GGO	609(14patients)	46(14patients)	
Reticular(RET)	210(8patients)	66(9patients)	
Honeycomb(HCM)	282(10patients)	73(9patients)	
Emphysema(EMP)	4406(10patients)	296(11patients)	
Nodular(NOD)	762 (9 patients)	65(10patients)	
Normal(NOR)	5371(11patients)	355(11patients)	
Total	11783	927	

大阪大学医学部病院から提供

ROI サイズ = 32x32 [pixels] ~ 16×16 [mm²]

頑張ってROIサンプルを取り出しても 1.1×10⁴個程度

DCNN モデルのパラメータ数と訓練画像

	LeCun+95	Fukushima+98 Fukushima+13	Krizhevsky+12	Simonyan+15	Ours [Krizhevsky+12]
Data type	Handwritten digits (MNIST)	Handwritten digits (ETL-1)	Natural images (ImageNet)	Natural images (ImageNet)	DLD ROI
DB size	6×10 ⁴	1.0×10 ⁴	1.2×10 ⁶	1.2×10 ⁶	1.1×10 ⁴
#layers	6	13	13	19	13
#weights	1.0×10 ⁵	6.9×10 ⁵	1.3×10 ⁷	1.4×10 ⁷	1.3×10 ⁷

一般に NN を訓練するには十分な数の学習データが必要

モデルとしての複雑度が高い 自由パラメータ 10⁷ オーダー

DNNを十分に学習させるには自由パラメータの1/10 の学習データ必要[Bengio+14]

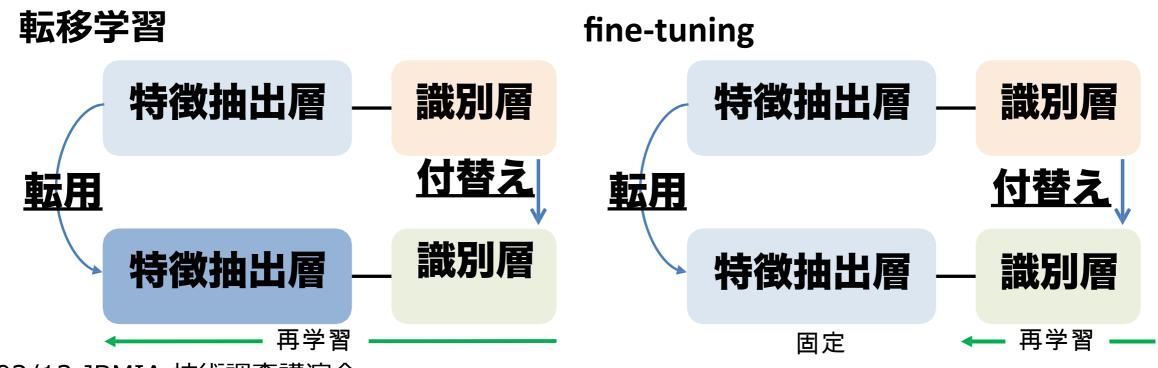
研究目標

- 学習データが少数かつテクスチャ構造を持つ ケースの問題に適用するための学習則の提案
 - □ 転移学習をベースにした逐次的な学習法
 - □ 入手が容易なデータにより特徴抽出部の性能向上を図る

- □ DCNNがテクスチャ画像に対する特徴表現を どのように学習するのかの解析
 - □ 学習過程による特徴表現の違いを可視化により分析
 - □ DCNNがテクスチャ認識で<mark>なぜ</mark>うまく働くのかを考察

転移学習

- ☑転移学習 (Transfer Learning)
 - ☑ 目的とは違う問題で得た知識を転用するテクニック
- ☑DCNNにおける転移学習
 - ▼特徴抽出の性能向上のため特徴抽出層を転用 識別層は新たに初期化しネットワーク全体を再学習
 - ☑ 識別層のみの再学習を行うfine-tuningとは異なる



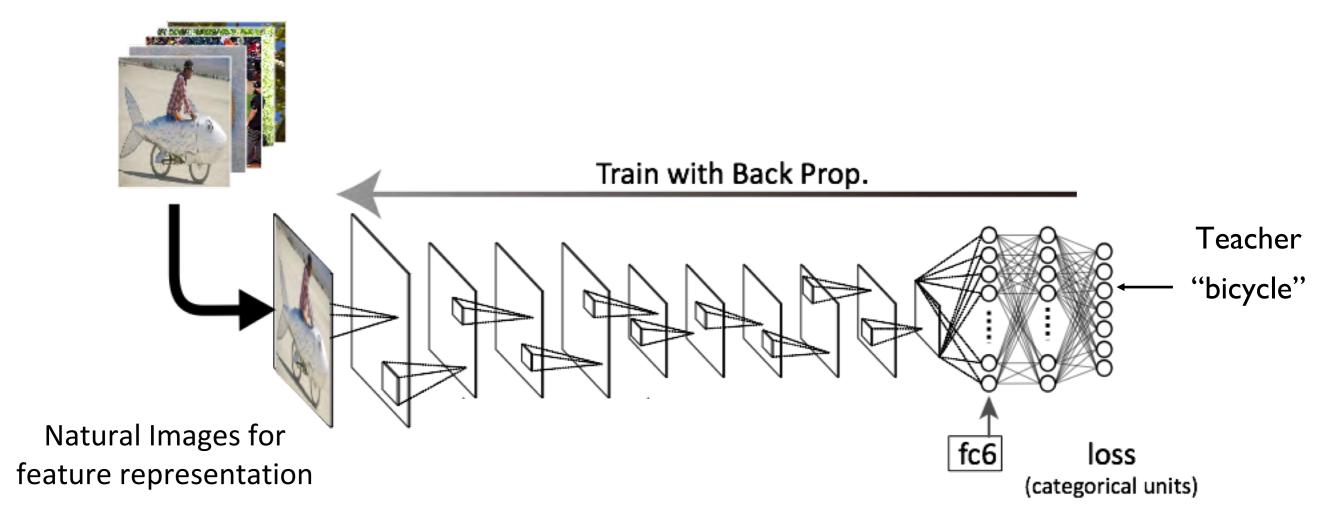
2019/02/12 JBMIA 技術調査講演会

小規模 DB で DCNN を学習させるには?

仮定したこと:[Shouno+15]

画像は何らかの共通構造を持っていて、DCNN はその構造を表現できてう

自然画像はDLDパターンの構造表現に役立つかも (特殊な画像は大人になってから学習)

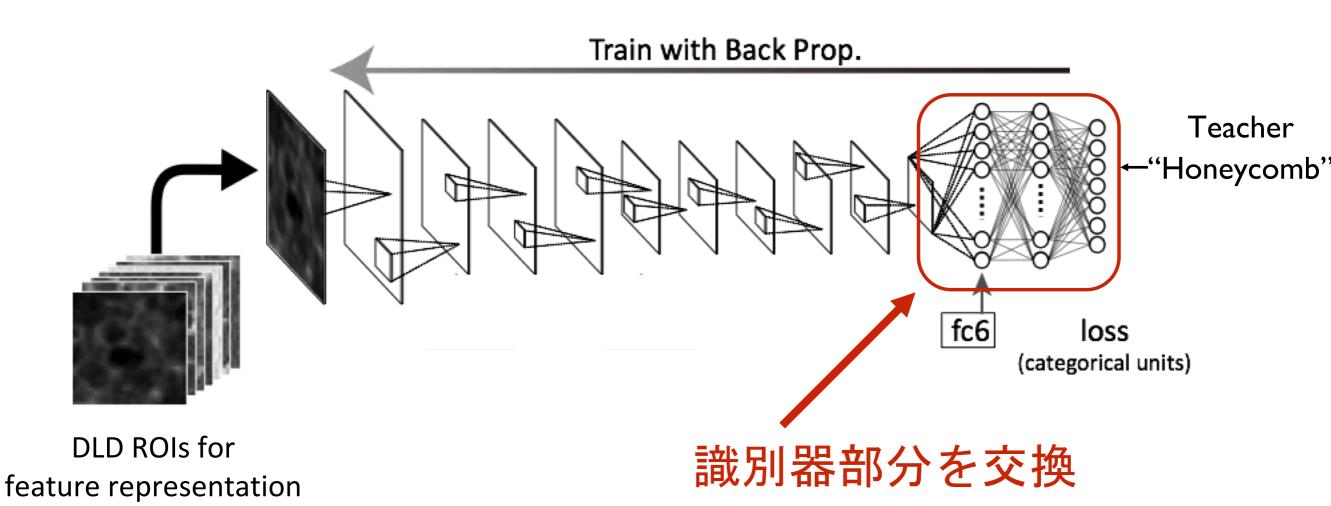


小規模 DB で DCNN を学習させるには?

仮定したこと:[Shouno+15]

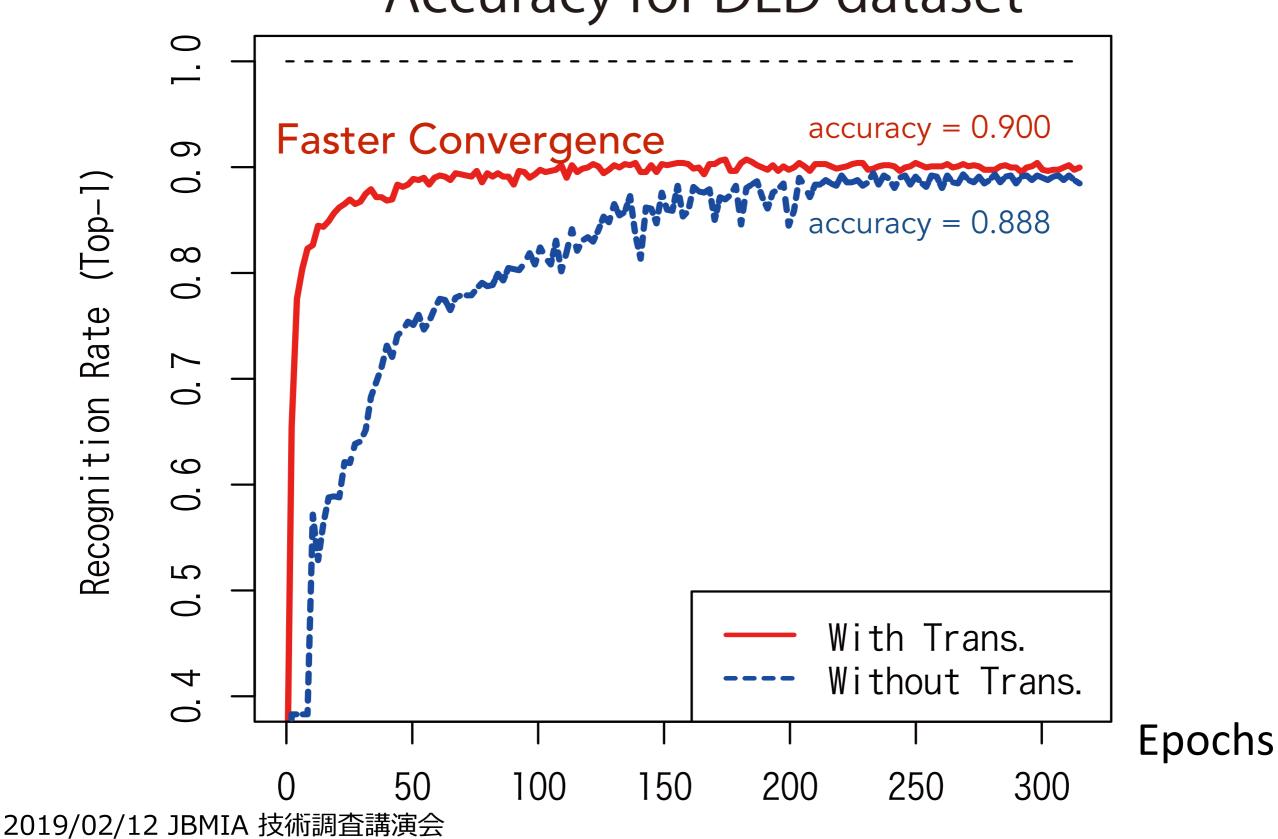
画像は何らかの共通構造を持っていて、DCNN はその構造を表現できる

自然画像はDLDパターンの構造表現に役立つかも (特殊な画像は大人になってから学習)



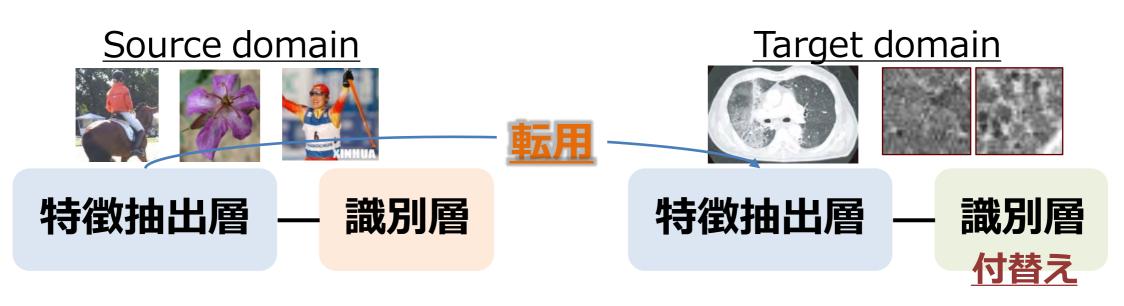
小規模 DB での DCNN 識別精度の推移





転移元は自然画像(だけ)でよいのか?

- · 自然画像を用いた転移学習 [Shouno+15], [Gao+16]
 - 自然画像をSource domainとした転移学習によって びまん性肺疾患陰影の識別を行うもの
 - 転移学習を行わない場合に比べて、DCNNの識別性能が 向上することが報告されている



Source domainとTarget domainは大きく異なっており そのドメイン選択の適切性に疑問が残る [Azizpour+16]

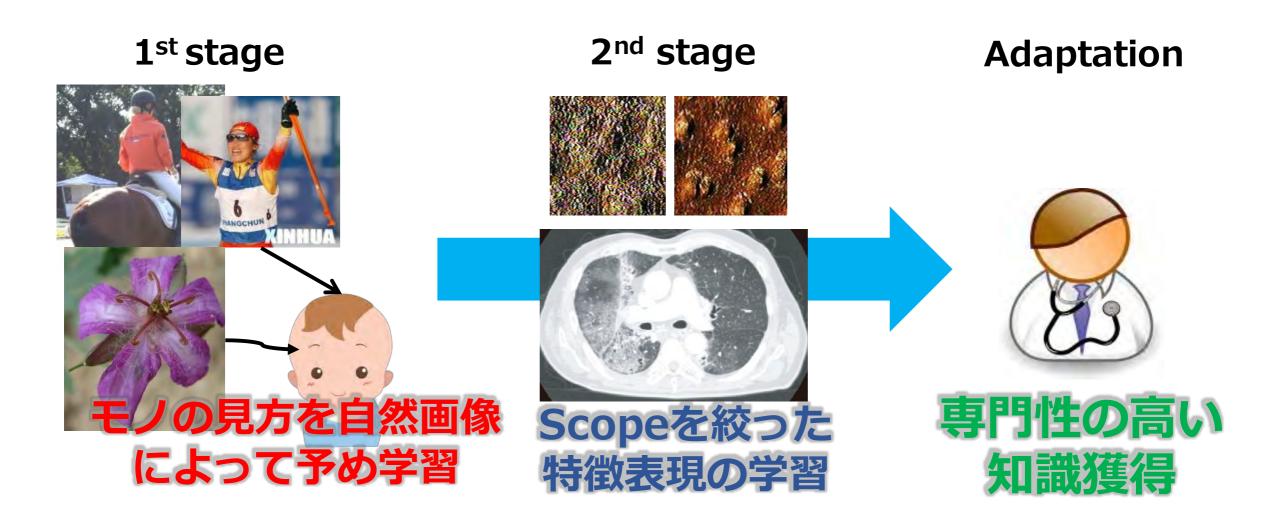
2段階転移学習への拡張(1)

自然画像とテスクチャ画像を用いた2段階に渡る転移学習

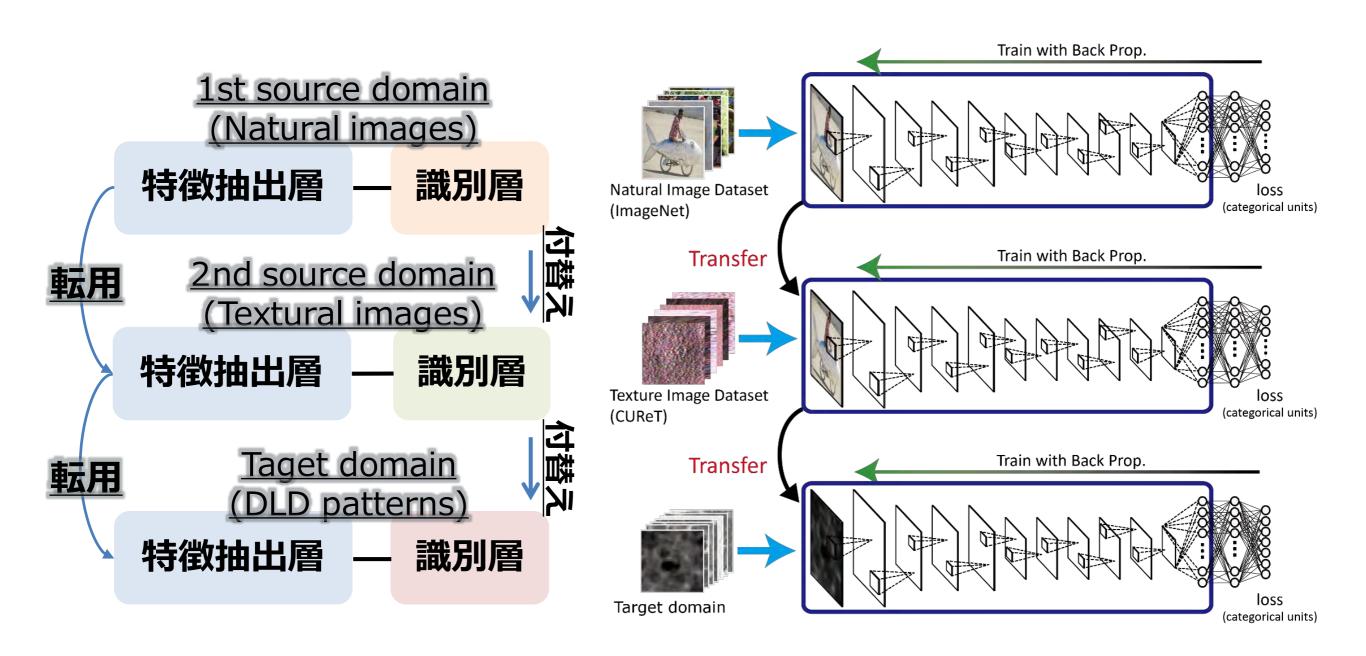
<u>目的</u>

(自然画像): 画像一般の特徴の獲得

(テクスチャ画像): テクスチャ画像に有用な特徴の獲得



2段階転移学習への拡張(2)



Source domainによる特徴抽出部の学習結果を初期値とし識別部を付け替えつつ逐次的に学習を行う

Source domain の選択

自然画像: ILSVRC2012 dataset

- 大規模一般画像データベースImageNetのサブセット
- 一般物体1000クラスの写真;計1,200,000枚



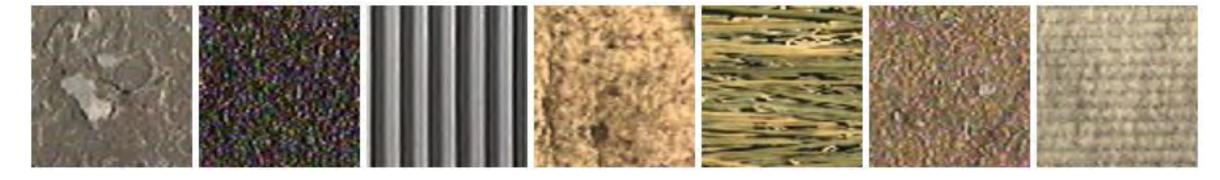






テクスチャ画像: CUReT database

- Utrecht Univ.が提供するテクスチャ画像データベース
- 61クラスの材料の拡大写真;約14,000枚 ⇒ **4倍に拡張**



比較するDCNNたち

学習過程の異なる4つのモデルの識別性能を比較する

- (1) 初期値からDLDのみで学習 **転移学習なし**
- (2) テクスチャ画像 → CT画像
- (3) 自然画像
- → CT画像

1段階転移学習

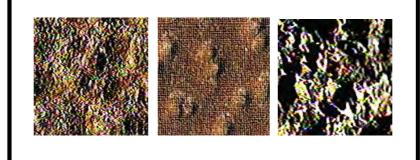
(4) 自然画像→テクスチャ画像→CT画像 **一 2段階転移学習** *(proposed)*

自然画像· ILSVRC2012



計 1,200,000枚

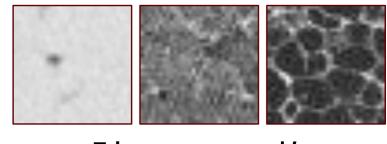
テクスチャ画像 CUReT texture DB



計 14,406枚

CT画像

肺断面の胸部X線CT <u>(※識別対象)</u>



計 12,888枚

比較方法

(1) 識別性能の比較

各学習の流れごとにモデルの識別性能を比較



転移学習がモデルの性能向上に寄与するかを検証

(2) 学習データ数に対する頑健性

学習データを人為的に減少させ、モデルの性能が どのように低下するかを検証・比較

■ 転移学習がモデルのデータ数に対する頑健性に どのように寄与するかを検証

考察: 学習済みモデルの特徴表現を比較し、 転移学習が「なぜ」うまく働くのかを考察

実験(1) における比較指標

それぞれのモデルについてDLDの識別性能を比較する

性能指標

• Accuracy =
$$\frac{TN + TP}{TN + FN + TP + FP}$$

• Recall =
$$\frac{TP}{TP + FN}$$

• Precision =
$$\frac{TP}{TP + FP}$$

• F1–Score =
$$\frac{2 \cdot Recall \cdot Precision}{Recall + Precision}$$

• Avg –
$$Loss = -\frac{1}{N} \sum_{y_i, t_i \in \mathfrak{D}}^{N} y_i \log t_i$$

TN: True Negative

TP: True Positive

FN: False Negative

FP: False Positive

y_i: DCNNの出力

 t_i : 真のラベル

- 大きいほど良
- 小さいほど良

2019/02/12 JBMIA 技術調査講演会

実験(1)の結果

	(1)	(2)	(3)	(4)
転移学習	なし	テクスチャ	自然画像	2段階 (proposed)
Accuracy	0.9277	0.9201	0.9558	0.9601
Recall	0.9583	0.9412	0.9484	0.9739
Precision	0.9590	0.9417	0.9471	0.9719
F1-Score	0.9583	0.9411	0.9470	0.9724

- ・2段階転移学習(4)は全ての性能指標で最も高い性能を示した
- ・テクスチャ画像のみを用いた転移学習(2)はどの性能指標も ランダム初期値より悪い性能を示した

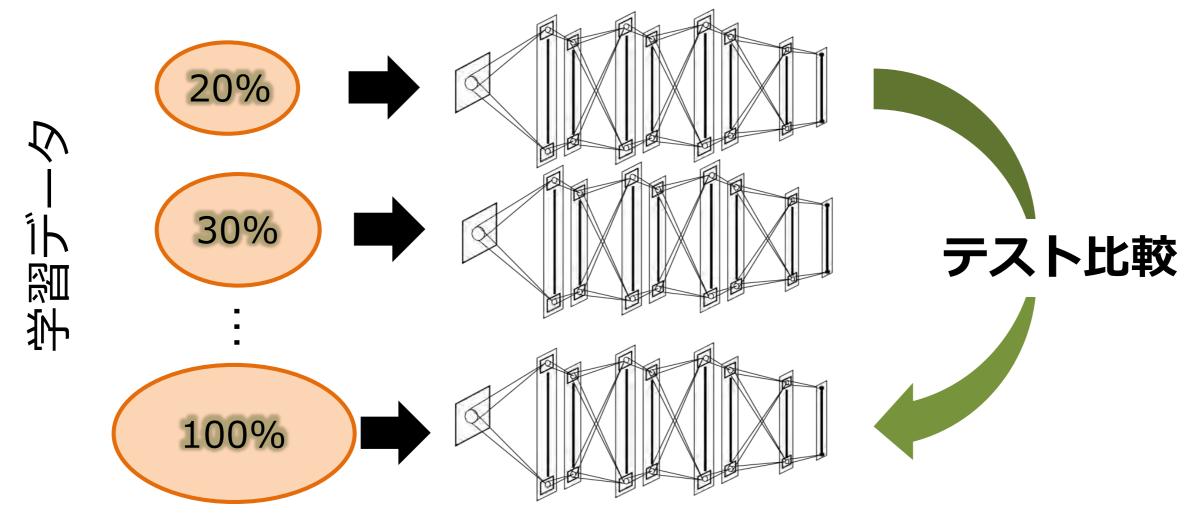


Source domain選択の重要性を示唆

実験(2): データ数に対する頑健性

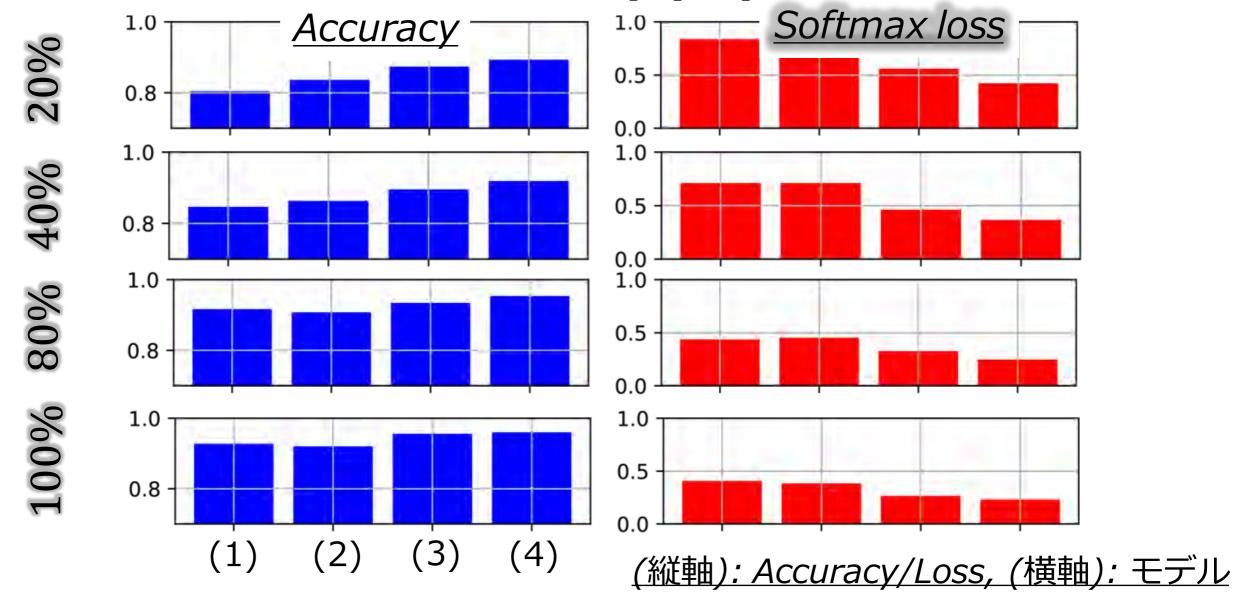
学習データ数の減少に対する頑健性を比較

- ・DLDの学習データ数を20%~100%に変化させた際の テストデータに対する性能の変化を(1)~(4)で観察
- ・転移学習によるモデル性能の頑健性の向上を検証



実験(2) の結果

同じ学習データを与えた場合の(1)~(4)のモデル性能を比較

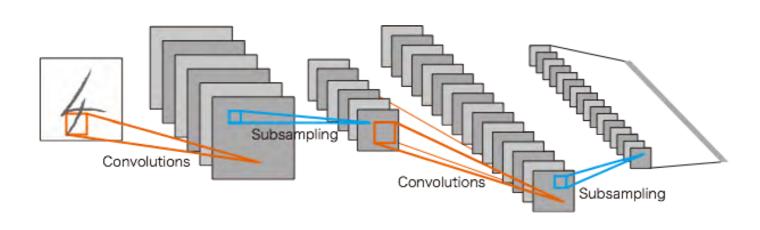


- 全ケースで2段階転移学習(4)は最も高い性能を示す
- 大きな減少比でScratch学習(1)とテクスチャ転移(2)が逆転(1) DLD のみ (2) CUReT→DLD (3) ImgNet→DLD (4) 2-staged(proposed)

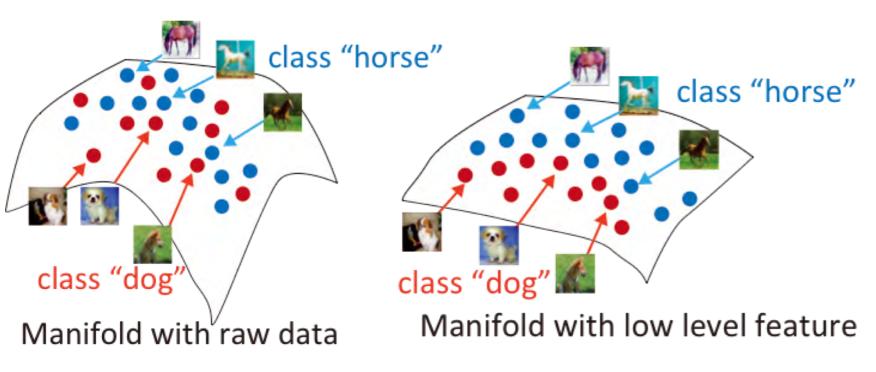
中間層の表現で 何がおこっているのか? 黒魔術の解明へ向けて

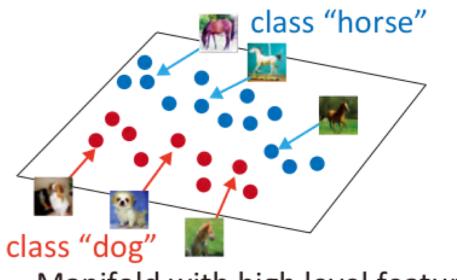
DCNN: 階層ごとの表現の違いは何か?

☑Deep Convolution net の本質は階層化 階層を経るごとに、どのような表現になるのか?



多様体仮説 [Refail 6, Bengio+16]





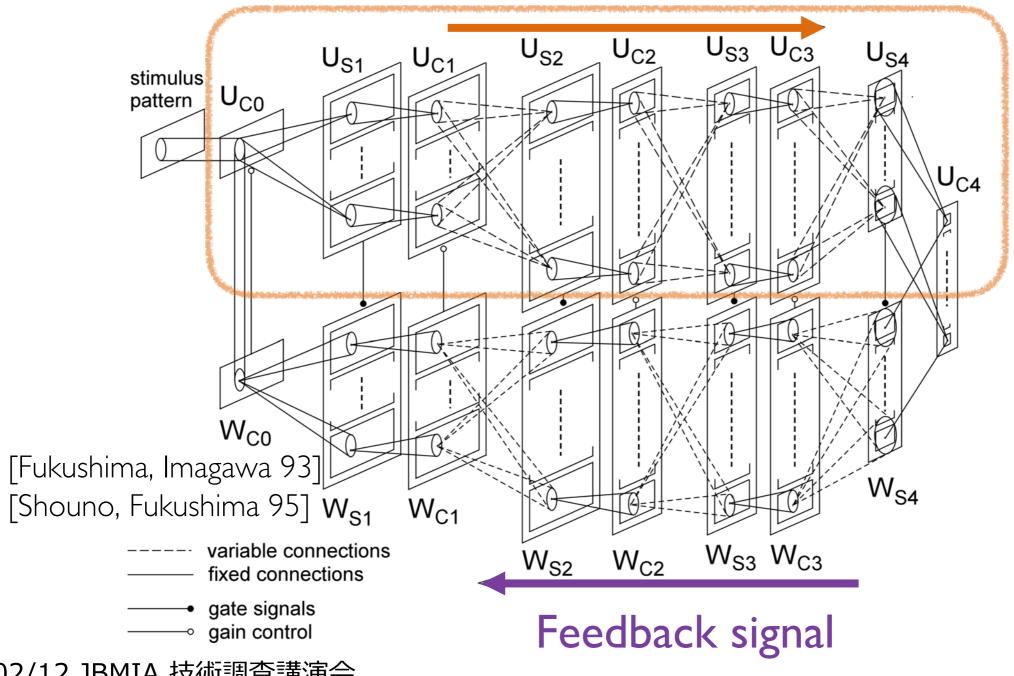
Manifold with high level feature

中間層の表現を直接見る

☑DCNNは {何処,何}を観ているのか?

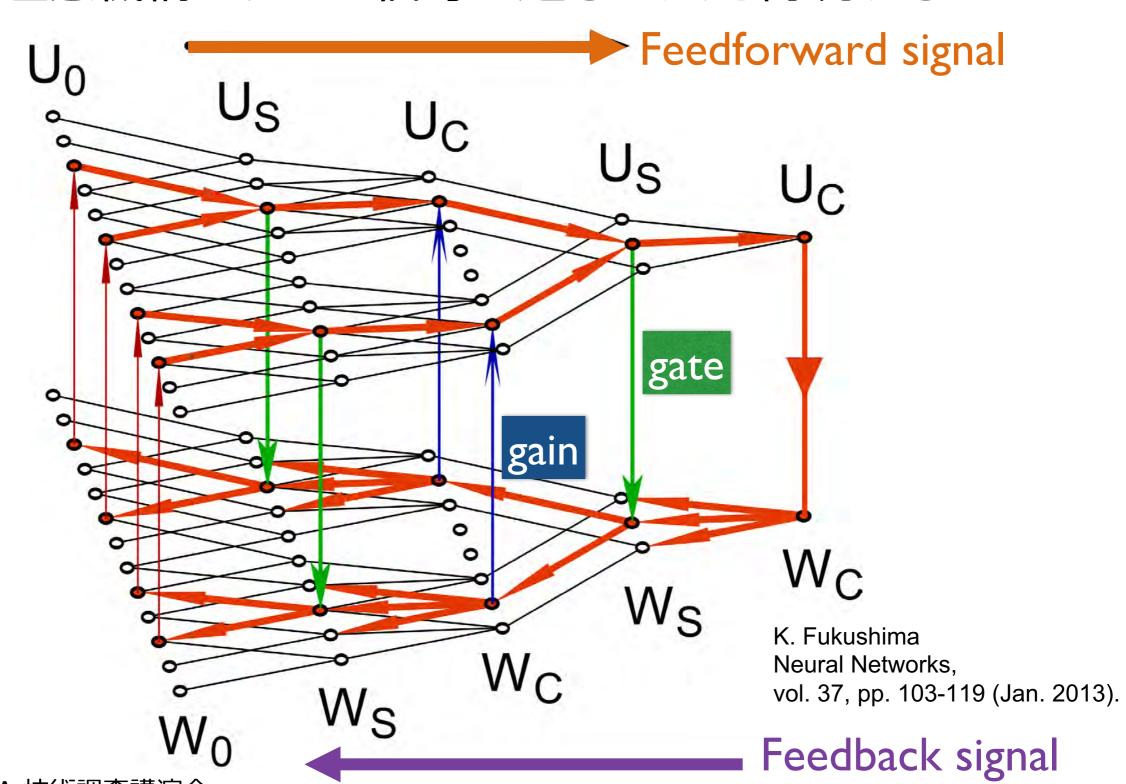
☑バックプロジェクション(選択的注意機構)の導入

[Fukushima 88, Shouno+95, Zeiler+14, 鈴木+17]

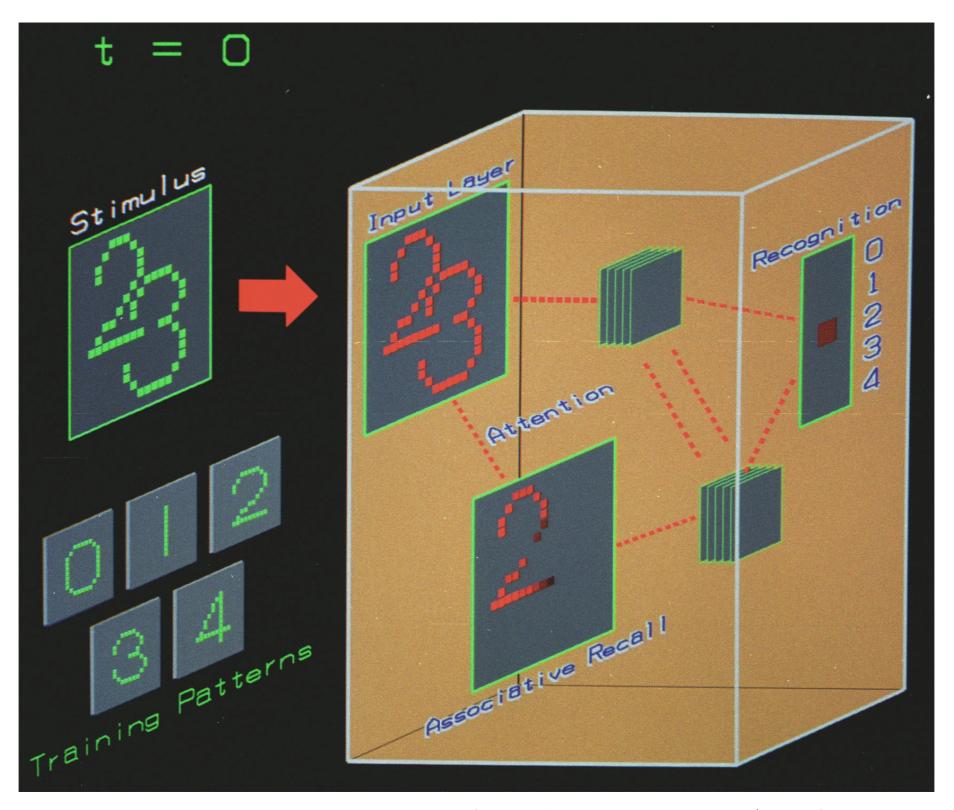


フィードバック信号の概略

選択的注意機構モデル: 信号が通るパスを再現する



選択的注意機構の動作イメージ



福島邦彦: "視覚パターン認識のモデル", 遺伝別冊2号, 「脳研究とニューロコンピュータ」, pp. 102-110 (1989.11).

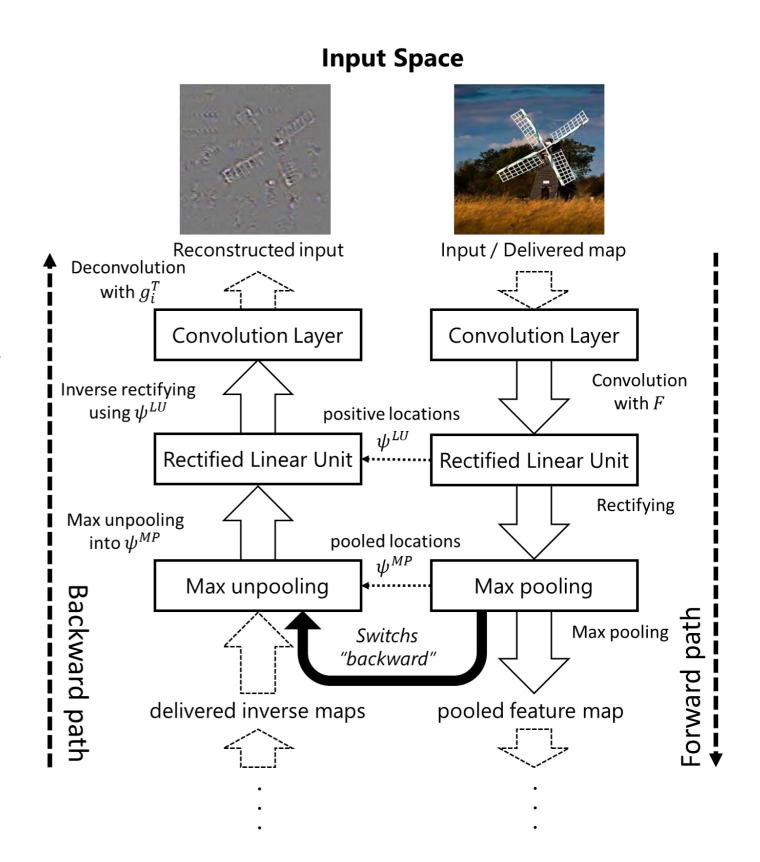
DeSaliNet

- ・[Mahendran+16]による 特徴表現の可視化手法
- ・入力の**どういった成分**が 特徴表現に寄与するかを 入力空間の画像として観察
- ・DCNNを一つの合成関数と 捉え、要素ごとの逆変換で 特徴表現を入力空間に復元

$$f = g_1 \circ g_2 \circ \cdots \circ g_n$$

$$\updownarrow$$

$$f^{\dagger} = g_n^{\dagger} \circ \cdots \circ g_2^{\dagger} \circ g_1^{\dagger}$$

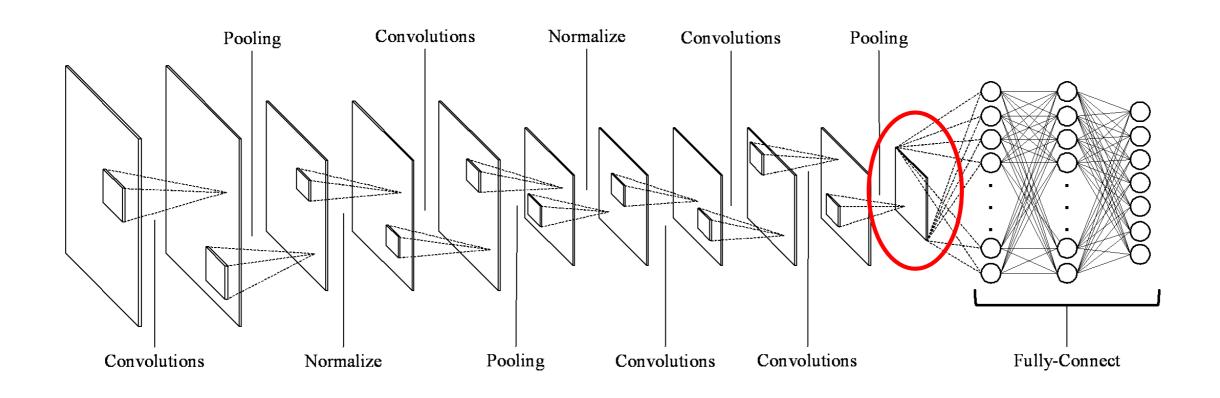


特徴の可視化

なぜ転移学習によって識別性能が向上するのか?



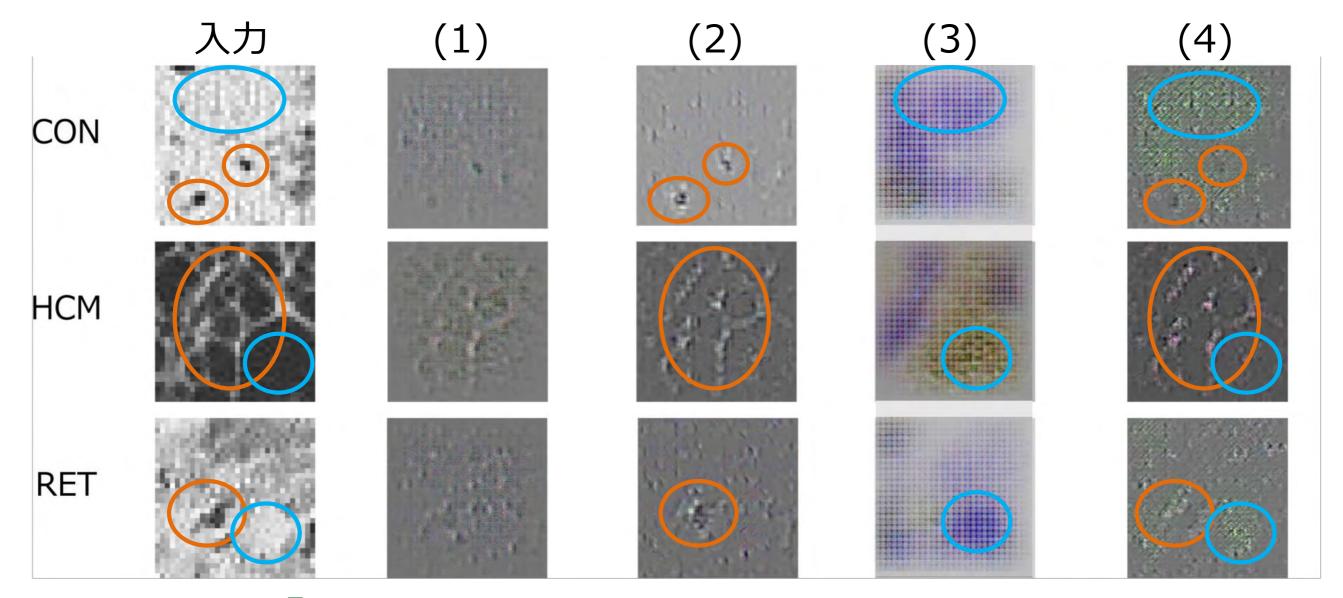
識別層に入力される特徴表現の変化に着目



転移元ドメインによる特徴表現の変化を観察することで DCNNにおける転移学習のメカニズムを理解する

特徴表現の可視化結果

- ・自然画像を学習したモデル(2)は入力のエッジ構造に テクスチャを学習したモデル(3)は模様に活性を示す
- ・2段階転移学習ではそれらの両方に活性がみられる



(1) DLD のみ (2) CUReT→DLD

(3) ImgNet→DLD (4) 2-staged(proposed)

まとめ

DCNNを少数のテクスチャデータセットの判別学習に適用するための2段階学習手法の提案

自然画像とテクスチャ画像のデータセットを用いた逐次的な 事前学習

びまん性肺疾患陰影の識別において、性能向上を確認できた DCNNの転移学習メカニズムについて特徴表現の可視化手法を 用いた分析

転移元ドメインに応じた特徴表現獲得を確認

転移元ドメイン選択の重要性が示唆された

禁無断転載

2018年度「ビジネス機器関連技術調査報告書」"I-2"部

発行 2019年6月

一般社団法人 ビジネス機械・情報システム産業協会 (JBMIA) 技術委員会 技術調査専門委員会

〒108-0073 東京都港区三田三丁目4番10号 リーラヒジリザカ7階

電話 03-6809-5010 (代表) / FAX 03-3451-1770